

QY	1232	CCAACTTCATCGAAGACCTCGAGAGATCCCTTGCTCTCCCGTGGGTGCTCCCTCATCT	1291
Db	2122	GCGGAATCATCGGGAUATTGGATACCTTGTGCTGACTCCGGTCAAGGTGGGTATGGTAT	2181
QY	1292	ATGGCGACCGGATACATCTGAACTGGTTCGGCGGTGAGCCGTTTTCCCTCGCTGCGA	1351
Db	2182	ATGGGGACCGGGAATATGCTTGTCCGTGGCGCGGGGAAGATCTCAGCCTGCTGGTG	2241
QY	1352	ACTACTCCCAAGCGCCGACAGTTCGGAAGCGCAGAGGTACACGCCCTGAAAGTCAACGGG	1411
Db	2242	AGTACGAGGATGCGGGAAGTTCGCTGCTGGGTATGCCAAGTGCAGACGAAATCAT	2301
QY	1412	TCAGATATGGG---AAACTGCGAGTATGGTAAATTTCTCTTCACTCGCTCTATGAGG	1468
Db	2302	CCTACGTTGGGGTCTAGTAAGGCAGTATGGAATCTTCGTTACGCGTGTCTTCAGG	2361
QY	1469	CAGGCCATGAAGTCCCATACTACAGCCCATCGCTCCCTGCAATGTTTAAACGGACTA	1528
Db	2362	CGGGCCATGAGTGCATTTTATCAGCCGAAACGGCGTATGAGATTTTAACTCGCGTC	2421
QY	1529	TCTTCGTTGGGATATCGCAGAGG	1553
Db	2422	AGTTTAATTGGGATATGCGACGG	2446
RESULT 11			
A75535			
LOCUS	A75535	2503 bp	DNA linear PAT 15-OCT-1999
DEFINITION	Sequence 2 from Patent WO9400579.		
ACCESSION	A75535		
VERSION	A75535.1	GI:6065609	
KEYWORDS			
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 2503)		
AUTHORS	Fleer R. and Fournier, A.		
TITLE	MODIFIED KLUTVEROMYCES YEASTS, THEIR PREPARATION AND USE		
JOURNAL	Patent: WO 9400579-A 2 06-JAN-1994;		
FEATURES	RHONE POULENC RORER SA (FR): FLEER REINHARD (FR)		
source	Location/Qualifiers		
gene	1..2503		
CDS	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/strain="Kluyveromyces lactis"		
	/db_xref="taxon:32644"		
	387..1862		
	/gene="kl.PRC1"		
	387..1862		
	/gene="kl.PRC1"		
	/codon_start=1		
	/product="Gene de la protease C de K.lactis"		
	/protein_id="CAB58521.1"		
	/db_xref="GI:6065610"		
	/translation="MVSIFELLISYGLWSVLAIISLNAVDSLFNSFDGNNIEDHE		
	TANYTQVFSFNIDDAISLRKLPDKSLGVDTVKQSGYLDYQSKHFFYWFPE		
	RNDPNDPVLWINGCGSGFVGLFELGPGSSIGADLKPTINPYNSNSASVFLDQ		
	PVGVSFGYDGSVTTDDAAKDVIFLDLFFERPHI-RNNDLHISGSYAGHYLFKIA		
	HEIAVVAEDSSNLSVSLINGTDPLTQYQYEPMACGEGGIPAVLEPEDCLDMNR		
	NLPCLSLVDRCVYKSHSVGLADRYCEQGITGVYEKGRNPYDIRSKCEAEDSGA		
	CYQEETIYSDYLNQEVQRALGTVDSFQCSSDVGIGFATGDPGSPFHQYVABLLD		
	QDINLVLAGDKDTCINLGNLAWTEKLEWRVNEYKQVLRVWKSEETDITGETKS		
	YGLTVLIRYLDAGHMVPHDPQENSLQWNSWIGNIAKRSRI"		
BASE COUNT	703 a	475 c	517 g 808 t
ORIGIN			
Query Match	5.2%;	Score 86.4;	DB 6; Length 2503;
Best Local Similarity	57.4%;	Pred. No. 3.7e-11;	
Matches 156;	Conservative	0;	Mismatches 116; Indels 0; Gaps 0;
QY	199	GTCAATCTCTCTGGATATGTGCACACTCTCCCGAGTCCCATACCTTCTTCTGGTTC	258
Db	624	GTGAACAATGTCGGGATATTAGATTACAGGACTCAAAACACTTCTTTATTGGTTT	683

REFERENCE 1 (bases 1 to 2002)
AUTHORS Xaver, D.S. and Thompson, S.A.
TITLE Gene encoding carboxypeptidase of aspergillus niger
JOURNAL Patent: US 5594119-A 3 14-JAN-1997;
FEATURES Location/Qualifiers
1..2002
source /organism="unknown"
BASE COUNT 416 a 591 c 505 g 490 t
ORIGIN

Query Match 5.0%; Score 84; DB 6; Length 2002;
Best Local Similarity 55.5%; Pred. No. 1.5e-10;
Matches 161; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

223 GACACCTCTCCCGAGTCCCATACCTTCTCTGTTTCTTCGAAGCCAGACATAAACCAGAA 282
|||
617 GACAAACGAGAACGACAAAGCATCTGTTCTCTGAGTCTCGCAATGACCCGAG 676
|||
283 ACTGCACCTATCACATTGTGTTGAATGGTGGCCCTCGAAGCGGATTTTGTATCGGTCTC 342
|||
677 AATGACCTGTGTTCTGTGCTGAACGGTGGCCCTGATCTCTTCCCTCACCGGTTT 736
|||
343 TTCGAAGAGTGGGCCCTTGCATCTCAATTCGACTTTTGATGACTATCATCAACCCCTCAC 402
|||
737 TTCATGGAGTCCGCCCTTAGCAGCATCAACAAGAGATCCAGCCGGTCTACAACGACTAC 796
|||
403 TCGTGGACGAGGCTCCCAATTTACTATTCTCTCCAGCCATTGGGAGTGGGTTTCA 462
|||
797 GCTTGGAACTCAACGGCTCGTGTCTCTCTTGACGAGCTGTCAACGTCGGTTACTCT 856
|||
463 TATAGTGTAGCGTTGTGATGGGTCCATTAACCTGTAACCTGGGTGCTCGA 512
|||
857 TACAGCACTCTGCTGTGTCAGCGACACCCGTTGCTGTGGGAGGAGCTCTA 906
|||

RESULT 14
174375 174375 2002 bp DNA linear PAT 03-APR-1998
LOCUS Sequence 3 from patent US 5688663.
DEFINITION I74375
ACCESSION I74375.1 GI:3010516
VERSION I74375.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2002)
AUTHORS Xaver, D.Sue. and Thompson, S. Ann.
TITLE Gene encoding carboxypeptidase of Aspergillus niger
JOURNAL Patent: US 5688663-A 3 18-NOV-1997;
FEATURES Location/Qualifiers
1..2002
source /organism="unknown"
BASE COUNT 416 a 591 c 505 g 490 t
ORIGIN

Query Match 5.0%; Score 84; DB 6; Length 2002;
Best Local Similarity 55.5%; Pred. No. 1.5e-10;
Matches 161; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

223 GACACCTCTCCCGAGTCCCATACCTTCTCTGTTTCTTCGAAGCCAGACATAAACCAGAA 282
|||
617 GACAAACGAGAACGACAAAGCATCTGTTCTCTGAGTCTCGCAATGACCCGAG 676
|||
283 ACTGCACCTATCACATTGTGTTGAATGGTGGCCCTCGAAGCGGATTTTGTATCGGTCTC 342
|||
677 AATGACCTGTGTTCTGTGCTGAACGGTGGCCCTGATCTCTTCCCTCACCGGTTT 736
|||
343 TTCGAAGAGTGGGCCCTTGCATCTCAATTCGACTTTTGATGACTATCATCAACCCCTCAC 402
|||
737 TTCATGGAGTCCGCCCTTAGCAGCATCAACAAGAGATCCAGCCGGTCTACAACGACTAC 796
|||
403 TCGTGGACGAGGCTCCCAATTTACTATTCTCTCCAGCCATTGGGAGTGGGTTTCA 462
|||

259 TTCGAAGCCAGACATACCAGAACTGCACCTATCACATTGTGTTGAATGTTGTCCT 318
|||
684 TTTGAGTCTAGAAATGACCCAGAGATGACCCGATGATGTTTAAACGGTGTCT 743
|||
319 GGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTGTGGGCCCTTGCATGTCGACT 378
|||
744 GGCTGTCTCTTTGCTGGTCTTTCTTTGAATGGGACCTTCTTCTATAGAGCTGAT 803
|||
379 TTTGATGACTACATCAACCCCTCACTCGTGGAGCGAGTCTCCAAATTTACTATTCCTGTC 438
|||
804 TTGAACCCATTATACCCCTACTCTTGGAAATCCAAACGTTCTGTGATATTCCTAGAT 863
|||
439 CAGCCATTGGGAGTCCGCTTTTTCATATAGTGA 470
|||
864 CAGCCTGTGGTGTGGGTTCTCATACGGTGA 895
|||

RESULT 12
170282 170282 2503 bp DNA linear PAT 03-APR-1998
LOCUS Sequence 3 from patent US 5679544.
DEFINITION I70282
ACCESSION I70282
VERSION I70282.1 GI:3006417
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2503)
AUTHORS Fleer, R., Fournier, A. and Yeh, P.
TITLE Modified Kluyveromyces yeasts, their preparation and use
JOURNAL Patent: US 5679544-A 3 21-OCT-1997;
FEATURES Location/Qualifiers
1..2503
source /organism="unknown"
BASE COUNT 703 a 475 c 517 g 808 t
ORIGIN

Query Match 5.2%; Score 86.4; DB 6; Length 2503;
Best Local Similarity 57.4%; Pred. No. 3.7e-11;
Matches 156; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

199 GTCAATCTCTACTCTGATATGTCGACACCTCTCCGAGTCCCATACCTTCTCTGTTTC 258
|||
624 GTCAACATGTCGGGATATTTAGATTACAGGACTCAAAACACTTCTTTATTGTTT 693
|||
259 TTCGAAGCCAGACATACCAGAACTGCACCTATCACATTGTGTTGAATGTTGTCCT 318
|||
684 TTTGAGTCTAGAAATGACCCAGAGATGACCCAGTATGATATGGTTAAACGGTGTCT 743
|||
319 GGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTGTGGGCCCTTGCATGTCGACT 378
|||
744 GGCTGTCTCTTTGCTGGTCTTTCTTTGAATGGGACCTTCTTCTATAGAGCTGAT 803
|||
379 TTTGATGACTACATCAACCCCTCACTCGTGGAGCGAGTCTCCAAATTTACTATTCCTGTC 438
|||
804 TTGAACCCATTATACCCCTACTCTTGGAAATCCAAACGTTCTGTGATATTCCTAGAT 863
|||
439 CAGCCATTGGGAGTCCGCTTTTTCATATAGTGA 470
|||
864 CAGCCTGTGGTGTGGGTTCTCATACGGTGA 895
|||

RESULT 13
I33983 I33983 2002 bp DNA linear PAT 06-FEB-1997
LOCUS Sequence 3 from patent US 5594119.
DEFINITION I33983
ACCESSION I33983
VERSION I33983.1 GI:1824774
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

Db 797 GCTTGAACCTCAACGGTCCGTGATCTTCCTTTACAGACGCTGTCAACGTCGGTTACTCT 856
QY 463 TATAGTGTATACGGTTGATGGTCCATTAAACCTGTAACTGGGTGCTGCGA 512
Db 857 TACAGCAACTCTGCTGTACAGACACCGTTGCTGCTGGCAAGACGCTCTA 906

RESULT 15
I77239
LOCUS I77239 2002 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 3 from patent US 5693510.
ACCESSION I77239
VERSION I77239.1 GI:3013393
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2002)
AUTHORS Xaver, D. Sue. and Thompson, S. Ann.
TITLE Gene encoding carboxypeptidase of Aspergillus niger
JOURNAL Patent: US 5693510-A 3 02-DEC-1997;
FEATURES Location/Qualifiers
source 1..2002
/organism="unknown"

BASE COUNT 416 a 591 c 505 g 490 t
ORIGIN

Query Match 5.0%; Score 84; DB 6; Length 2002;
Best Local Similarity 55.5%; Pred. No. 1.5e-10;
Matches 161; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 223 GACACCTCTCCGAGTCCCATACCTTCTTGGTTCTTCGAAGCCAGACATAACCCAGAA 282
Db 617 GACAAAGAGACGACCAAGCATCTGTCTACTGGTTCTTCGACTCTCGCAATGACCCCGAG 676
QY 283 ACTGCACCTATCACATTGTTGTTGAATGGTGGCCCTGGAAGCGATTCTTTGATCGGTCTC 342
Db 677 AATGACCTGTGTTCTGTGGCTGAACGGTGGCCCTGGATGCTTTCCTCACCCTGCTT 736
QY 343 TTCGAAGAGTTGGGCCCTTGGCCATGTCAATTGACTTTTGTGACTACATCAACCCCTCAC 402
Db 737 TTCATGGAGCTCGGCCCTAGCAGCATCAACAGAAGATCCAGCCGGTCTACAACGACTAC 796
QY 403 TCGTGAACGAGGTCCTCAATTACTATTCCTGTCGCCAGCCATTGGGAGTCGGCTTTTCA 462
Db 797 GCTTGGAACTCCACGCGCTCGGTGATCTTCCTTGACCAAGCTGTCACGTCGGTTACTCT 856
QY 463 TATAGTGTATACGGTTGATGGTCCATTAAACCTGTAACTGGGTGCTGCGA 512
Db 857 TACAGCAACTCTGCTGTACAGCACACCGTTGCTGCTGGCAAGGAGCTCTA 906

Search completed: November 21, 2003, 13:23:24
Job time : 6074 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:45:55 ; Search time 74.5 Seconds
(without alignments)
11555.215 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 555
Sequence: 1 atgcgtggctacgaattctt.....ccagtgtgtgcatgcatag 1668

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1660916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xl
-Q/cgn2_1/USPTO.spool_p/US09712338/runat 17112003 170226 10771/app query.fasta 1.1863
-DB=SPTRMBL 23 -QFMT=fastan -SUFFIX=oli.rspt -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1 1 127 @runat 17112003 170226 10771 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SPTRMBL 23:

- 1: sp archaea.*
- 2: sp bacteria.*
- 3: sp fungi.*
- 4: sp human.*
- 5: sp invertebrate.*
- 6: sp mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp page.*
- 10: sp plant.*
- 11: sp rodent.*
- 12: sp_virus.*
- 13: sp vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	454	81.8	555	3 Q96VZ9	Q96vz9 aspergillus

2	10	1.8	460	3	Q12569
3	9	1.6	67	2	Q937G9
4	9	1.6	145	16	Q8RDZ1
5	5	1.6	147	2	Q936V3
6	9	1.6	164	5	Q9GRR4
7	9	1.6	290	16	Q8B686
8	9	1.6	375	17	Q8TU72
9	9	1.6	405	2	P66024
10	9	1.6	405	2	P71195
11	9	1.6	405	2	Q91UM7
12	9	1.6	405	2	Q56451
13	9	1.6	405	2	Q9F3W1
14	9	1.6	468	10	Q9SHU3
15	9	1.6	469	10	Q9FH06
16	9	1.6	473	10	Q9FH05
17	9	1.6	479	10	Q9WAR8
18	9	1.6	495	2	Q48384
19	9	1.6	1535	5	Q9VFF0
20	9	1.6	2105	5	Q17679
21	8	1.4	51	8	Q957T5
22	8	1.4	63	8	Q957T1
23	8	1.4	63	8	Q957S9
24	8	1.4	76	4	Q96GB0
25	8	1.4	91	16	Q99R48
26	8	1.4	91	16	Q8NUN8
27	8	1.4	99	17	Q971Q8
28	8	1.4	114	17	Q8TJ61
29	8	1.4	116	15	Q75691
30	8	1.4	166	17	O59236
31	8	1.5	169	10	Q9ASH5
32	8	1.5	172	16	Q8XX16
33	8	1.5	184	17	Q27824
34	8	1.5	191	10	Q8LJM1
35	8	1.5	203	16	Q8DEH0
36	8	1.4	210	17	Q9HMH2
37	8	1.5	249	4	Q8H8T6
38	8	1.4	259	2	Q9S3T5
39	8	1.4	263	2	Q9ZF75
40	8	1.4	264	10	Q9LXY6
41	8	1.4	264	10	Q8GVD2
42	8	1.4	270	16	Q9RRC8
43	8	1.4	281	16	Q9K332
44	8	1.5	290	16	Q8ZGZ9
45	8	1.5	326	2	O68878

ALIGNMENTS

RESULT 1

Q96VZ9
ID Q96VZ9 PRELIMINARY; PRT; 555 AA.
AC Q96VZ9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Carboxypeptidase S1 precursor (EC 3.4.16.6).
GN CPl.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OC NCBI_TaxID=5062;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK3;
RA van den Broek P.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF394242; AAK77166.1; -
DR InterPro: IPR001563; Serine carboxypeptidase.
DR Pfam: PF00450; serine carboxypeptidase; 1.
DR PRINTS: PR00724; CRBOXYPTASEC.
DR ProDom: PD001189; Serine carboxypeptidase; 1.
DR PROSITE: PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase; Hydrolase; Signal.

QY	901	TACATGAAGTTTGCCCAACCAATGCCAAATGGTTGCCAGATTGTTGATTCACCTGCAAA	960
DB	301	TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys	320
QY	961	CAGACAAACCGCACCGCATTAGCTAGCTAGCCCTCTGCGCGAAGCCACCAATGTGC	1020
DB	321	GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys	340
QY	1021	AGGACAATGTTGAGGGCCATACCTACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
DB	341	ArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArg	360
QY	1081	CATCCATATGATGACCGGACTCCGCCAAGTTATTACAAATAATTTCTGCGAAAGGACTCT	1140
DB	361	HisProTyrAspAspProThrProSerTyrTyrAsnLysPheLeuAlaLysAspSer	380
QY	1141	GTCAATGAGCGCTATCGCGCTCAACATCACTACACCCAGTCCCAATAATGACGTCTACT	1200
DB	381	ValMetAspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyr	400
QY	1201	GCTTTCAGCAACAGCGGACTTGTCTGCGCCCACTTCATCGAAGACCTCGAGGAGATC	1260
DB	401	AlaPheGlnGlnThrGlyAspPheValTyrProAsnPheIleGluAspLeuGluGluIle	420
QY	1261	CTTGCTCTCCCGCTGCTCTCCCTCATCTATGCGCGACGCGGATTATCATCTGCAACTGG	1320
DB	421	LeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTyr	440
QY	1321	TTCGCGGCTCAGGCGGCTTTCCTCTGCTGCGAATCTCTCCCAAGCGCCGCTTCCGAGC	1380
DB	441	PheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSer	460
QY	1381	GCAGGTACAGCCCTGAAAGTCAACGGGCTCGAGTATGGGAACTCGCGAGTATGCT	1440
DB	461	AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyIleThrArgGluTyrGly	480
QY	1441	AATTTCTCTTCACTCGCGTCTATGAGCAGGCGCATGAAGTCCCATCTACACAGCCCATC	1500
DB	481	AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle	500
QY	1501	GCTTCCTTCACTGCGCTTAAACCGGACTATCTTCGTTGGGATATCGCAGGCGCCAGAG	1560
DB	501	AlaSerLeuGlnLeuPheAsnArgThrIlePheGlyTyrAspIleAlaGluGlyGlnLys	520
QY	1561	AAGATCTGCCCGCAGCTACAGACGAATGAGCGCTACAGCTACGATACGATACACAGTCT	1620
DB	521	LysIleTyrProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer	540
QY	1621	GTGCGCTGCTTACCGCTTACCGCATGTCCAGTGTGTGTATGCGCA	1685
DB	541	ValProLeuProThrAlaThrSerMetSerValGlyMetAla	555
RESULT 2			
ID	Q12569	PRELIMINARY;	PRT; 460 AA.
AC	Q12569	(1-EMBLrel. 01, Created)	
DT	01-NOV-1996	(1-EMBLrel. 01, Last sequence update)	
DT	01-NOV-1996	(1-EMBLrel. 01, Last sequence update)	
DT	01-JUN-2002	(1-EMBLrel. 21, Last annotation update)	
DE	Prepro-carboxypeptidase Z (EC 3.4.16.5).		
GN	SPCZ.		
OS	Absidia zychae.		
OC	Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;		
OC	Absidia.		
CO	NCBI_TaxID=36079;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NEIC 1199;		
RX	MEDLINE=95308538; PubMed=7789719;		
RA	Lee B.R., Takeuchi M., Kobayashi Y.;		
RT	"Molecular cloning and sequence analysis of the scpZ gene encoding the		
RT	serine carboxypeptidase of Absidia zychae.";		
RL	Curr. Genet. 27:159-165 (1995).		

DR EMBL; D16519; BAA03966.1; -;
 DR HSSP; P00729; IYSC.
 DR MEROPS; S10.010; -;
 DR InterPro; IPR001563; Serine carboxypept.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF00450; serine carboxypept; 1.
 DR PRINTS; PR00724; CRBOXYPASEC.
 DR ProDom; PD001189; Serine carboxypept; 1.
 KW Carboxypeptidase; Hydrolase.
 FT CHAIN 52 460 MATURE CARBOXYPEPTIDASE Z.
 SQ SEQUENCE 460 AA; 50825 MW; CE3A410764548BB6 CRC64;

Alignment Scores:
 Pred. No.: 0.654 Length: 460
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.80% Indels: 0
 DB: 3 Gaps: 0

US-09-712-338-1 (1-1668) x Q12569 (1-460)

QY 1456 CCGCTCTATGAGCGAGGCCATGAAGTCCCA 1485

Db 430 ArgValTyGluAlaGlyHisGluValPro 439

RESULT 3

ID Q937G9 PRELIMINARY; PRT; 67 AA.
 AC Q937G9;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Putative transposition protein (Fragment).
 GN TNUQ.
 OS Klebsiella sp. L513-39.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=143776;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=LS13-39; TRANSPOS=TN5056;
 RX MEDLINE=21604134; PubMed=11763242;
 RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
 RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
 RA Yurieva O.V., Nikiforov V.G.;
 RT "Mercury resistance transposons of Gram-negative environmental
 bacteria and their classification."
 RL Res. Microbiol. 152:811-822(2001).
 DR EMBL; AJ302770; CAC82957.1; -;
 FT NON-TER 1 1
 SQ SEQUENCE 67 AA; 7547 MW; 3EECC0A1149C62D9 CRC64;

Alignment Scores:
 Pred. No.: 9.83 Length: 67
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q937G9 (1-67)

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288

Db 3 GluAlaArgHisAsnProGluThrAla 11

RESULT 4

ID Q8RDZ1 PRELIMINARY; PRT; 145 AA.
 AC Q8RDZ1;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Integral membrane protein.
 GN FNI350.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 OC Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fonstein M., Kyrpides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 nucleatum strain ATCC 25586."
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE016639; AAL95546.1; -;
 KW Complete proteome.
 SQ SEQUENCE 145 AA; 15902 MW; 7DB20993CC1C6F0E CRC64;

Alignment Scores:
 Pred. No.: 8.82 Length: 145
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 16 Gaps: 0

US-09-712-338-1 (1-1668) x Q8RDZ1 (1-145)

QY 761 TTAATGGTTCAGCTTAATTTCAACT 787

Db 131 LeuMetValPheSerLeuIleSerThr 139

RESULT 5

ID Q936V3 PRELIMINARY; PRT; 147 AA.
 AC Q936V3;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Putative transposition protein ThiQ (Fragment).
 GN TNUQ.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=BW20; TRANSPOS=TN5053v4;
 RX MEDLINE=21604134; PubMed=11763242;
 RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
 RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
 RA Yurieva O.V., Nikiforov V.G.;
 RT "Mercury resistance transposons of Gram-negative environmental
 bacteria and their classification."
 RL Res. Microbiol. 152:811-822(2001).
 DR EMBL; AJ296313; CAC82974.1; -;
 FT NON-TER 1 1
 SQ SEQUENCE 147 AA; 16500 MW; 24636FA7A1489ABD CRC64;

Alignment Scores:
 Pred. No.: 8.81 Length: 147
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q936V3 (1-147)

```

Alignment Scores:
Pred. No.: 8 Length: 290
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 16 Gaps: 0

US-09-712-338-1 (1-1668) x Q8D686 (1-290)
QY 227 CCTCTCCGAGTCCCATACCTCTTCT 253
Db 161 ProLeuProSerProLeuProSerSer 169

RESULT 8
Q8TU72 PRELIMINARY; PRT; 375 AA.
ID Q8TU72
AC Q8TU72:
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 3-isopropylmalate dehydrogenase.
GN LEUB OR MA0201.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OC NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA FitzHugh W., Calvo S., Stange-Thomann N., Dearrellan K., Johnson R.,
RA Allen N., Naylor J., McKernan K., Talanas J., Tirrell A., Ye W.,
RA Linton L., McEwan P., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Zimmer A., Barber R.D., Cann I., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reave J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010678; AAM03654.1; -.
DR InterPro; IPR001804; Isohd.
DR Pfam; PF00180; isodh; 1.
DR PROSITE; PS00470; IDH_IMDH; 1.
KW Complete proteome.
SQ SEQUENCE 375 AA; 41318 MW; 35646A9B4A005AFC CRC64;

Alignment Scores:
Pred. No.: 7.72 Length: 375
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.65% Indels: 0
DB: 17 Gaps: 0

US-09-712-338-1 (1-1668) x Q8TU72 (1-375)
QY 1272 GGGGAGAGCAGATCTCTCGAGGTC 1246
Db 147 GlycylserLeuAspLeuLeuGluVal 155

RESULT 9
Q66024 PRELIMINARY; PRT; 405 AA.
ID Q66024
AC Q66024:
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

```

```

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
Db 83 GluAlaArgHisAsnProGluThrAla 91

RESULT 6
Q9GRR4 PRELIMINARY; PRT; 164 AA.
ID Q9GRR4
AC Q9GRR4:
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Conserved hypothetical protein L5882.03.
GN L5882.03.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Friedlin;
RA Masuy D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL445943; CAC14321.1; -.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17935 MW; F60D849331340AE7 CRC64;

Alignment Scores:
Pred. No.: 8.67 Length: 164
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.65% Indels: 0
DB: 5 Gaps: 0

US-09-712-338-1 (1-1668) x Q9GRR4 (1-164)
QY 185 TCGGAGAGCCCTCTGCCCGGTTCC 159
Db 154 SerGlnThrProSerAlaProGlySer 162

RESULT 7
Q8D686 PRELIMINARY; PRT; 290 AA.
ID Q8D686
AC Q8D686:
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN VY20650.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OC NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016810; AAC07593.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 290 AA; 33267 MW; AC8FF330BE7F0562 CRC64;

```

DE TniQ.
GN TniQ.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OG Plasmid pPB.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OX;
RX MEDLINE=98148002; PubMed=9479042;
RA Reniero D., Mozzon E., Galli E., Barbieri P.;
RT "Two aberrant mercury resistance transposons in the Pseudomonas
stutzeri plasmid pPB.";
RL Gene 208:37-42(1998).
DR EMBL; U81032; AAC38239.1; -.
KW Plasmid.
SQ SEQUENCE 405 AA; 45761 MW; EBB533CCE6C7CFF CRC64;

Alignment Scores:
Pred. No.: 7.64 Length: 405
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x O65024 (1-405)

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 10
P71195 ID P71195 PRELIMINARY; PRT; 405 AA.
AC P71195;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Plasmid R751, complete sequence.
OS Enterobacter aerogenes (Aerobacter aerogenes).
OG Plasmid R751.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=548;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98428622; PubMed=9753548;
RA Thorsted P.B., Macartney D.P., Akhtar P., Haines A.S., Ali N.,
RA Davidson P., Stafford T., Pocklington M.J., Pansegrau W.,
RA Wilkins B.M., Lanka E., Thomas C.M.;
RT "Complete sequence of the IncPbeta plasmid R751: implications for
evolution and organisation of the IncP backbones.";
RL J. Mol. Biol. 282:969-990(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Thomas C.M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Haines A.S., Thomas C.M.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U67194; AAC64465.1; -.
KW Plasmid.
SQ SEQUENCE 405 AA; 45567 MW; 0F0749CDA19244F7 CRC64;

Alignment Scores:
Pred. No.: 7.64 Length: 405
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0

DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x P71195 (1-405)

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 11
Q91UM7 ID Q91UM7 PRELIMINARY; PRT; 405 AA.
AC Q91UM7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE TniQ protein.
GN TniQ.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSB102
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=mercury resistance transposon Tns718;
RA Schweiker S., Keller M., Droegge M., Lanka E., Puehler A.,
RA Selbitschka W.;
RT "The genetic organization and evolution of the broad-host-range
mercury resistance plasmid pSB102 isolated from a microbial population
residing in the rhizosphere of alfalfa.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ304453; CAC79210.1; -.
KW Plasmid.
SQ SEQUENCE 405 AA; 45711 MW; 2396F2C89A2DE7B2 CRC64;

Alignment Scores:
Pred. No.: 7.64 Length: 405
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q91UM7 (1-405)

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 12
Q56451 ID Q56451 PRELIMINARY; PRT; 405 AA.
AC Q56451;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE Mercury resistant transposon TNS053 (Central Asia).
GN TniQ.
OS Xanthomonas sp.
OG Plasmid RP1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=29446;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93253772; PubMed=8387603;
RA Kholodii G.Y.A., Yurieva O.V., Lomovskaya O.L., Gorlenko Zh.M.,
RA Mindlin S.Z., Nikiforov V.G.;
RT "Tns053, a mercury resistance transposon with integron's ends.";
RL J. Mol. Biol. 230:1103-1107(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96130850; PubMed=8594337;

Mon Nov 24 13:41:02 2003

us-09-712-338-1.oli.rspt

RA Kholodii G.Y., Mindlin S.Z., Bass I.A., Yurieva O.V., Minakhina S.V.,
 RA Nikiforov V.G.;
 RT "Four genes, two ends, and a res region are involved in transposition
 of Tn5053: a paradigm for a novel family of transposons carrying either
 RT a mer operon or an integron.";
 RL Mol. Microbiol. 17:1189-1200(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kholodii G.Y.;
 RL Russ. J. Genet. 31:1447-1451(1995).
 DR EMBL; L40585; AAA98331.1; -;
 KW Plasmid.
 SQ SEQUENCE 405 AA; 45830 MW; 89880020E32684C3 CRC64;

Alignment Scores:
 Pred. No.: 7.64 Length: 405
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q56451 (1-405)

OY 262 GAAGCCAGACATAACCCAGAACTGCA 288
 DB 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 13

ID Q9F3W1 PRELIMINARY; PRT; 405 AA.
 AC Q9F3W1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE TniQ protein.
 GN TniQ.
 OS Pseudomonas sp. ED23-33.
 OC Bacteria; Proteobacteria.
 OX NCBI_TaxID=83781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=Tn5058;
 RA Minakhina S., Minakhin L., Kholodii G., Mindlin S., Gorlenko Z.H.,
 RA Yurieva O., Nikiforov V.;
 RT "Molecular inventory of transposons from environmental bacteria:
 RT epidemic dissemination of Tn21-, Tn5041-, and Tn5053-related
 RT transposons.";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y17897; CAC14709.1; -;
 SQ SEQUENCE 405 AA; 45706 MW; 033A913AF44E43EE CRC64;

Alignment Scores:
 Pred. No.: 7.64 Length: 405
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q9F3W1 (1-405)

OY 262 GAAGCCAGACATAACCCAGAACTGCA 288
 DB 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 14

ID Q9SHU3 PRELIMINARY; PRT; 468 AA.
 AC Q9SHU3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative serine carboxypeptidase II.
 GN At2G12480.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=cv. Columbia;
 RC MEDLINE=20083487; PubMed=10617197;
 RX Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.W., Somerville C.R.,
 RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=cv. Columbia;
 RC Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007288; AAD28682.1; -;
 DR HSSP; P08819; 1WHT.
 DR MEROPS; S10.005; -;
 DR InterPro; IPR002000; Lamp.
 DR InterPro; IPR001563; Serine carboxypept.
 DR Pfam; PF00450; serine carboxypept. 1.
 DR PRINTS; PR00724; CRBOXYPRTASEC.
 DR ProDom; PD001189; Serine carboxypept. 1.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00310; LAMP_1; 1.
 KW Carboxypeptidase.
 SQ SEQUENCE 468 AA; 52836 MW; 47DD2055AB9AE1D7 CRC64;

Alignment Scores:
 Pred. No.: 7.48 Length: 468
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 10 Gaps: 0

US-09-712-338-1 (1-1668) x Q9SHU3 (1-468)

OY 295 ACATTTGGTGTGAATGGTGGCCCTGGA 321
 DB 77 ThrLeuTrpLeuAsnGlyClyProGly 85

RESULT 15

ID Q9FH06 PRELIMINARY; PRT; 469 AA.
 AC Q9FH06;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Serine carboxypeptidase-II like.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Columbia;
 RC MEDLINE=20181125; PubMed=10718197;
 RX Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence

RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC

clones.",
DNA Res.: 7:31-63(2000).

DR EMEL; AB023032; BAB10196.1; -.

DR HSP; P08819; 1WHT.

DR MEROPS; S10.005; -.

DR InterPro; IPR001563; Serine_carbpept.

DR Pfam; PF00450; serine_carbpept; 1.

DR PRINTS; PR00724; CREOXYPTASEC.

DR ProDom; PD001189; Serine_carbpept; 1.

DR PROSITE; PS00131; CARBOXYPEPT_SER_SBR; 1.

KW Carboxypeptidase.

SQ SEQUENCE 469 AA; 52541 MW; 90357E27B9B64527 CRC64;

Alignment Scores:

Pred. No.: 7.48 Length: 469

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.62% Indels: 0

DB: 10 Gaps: 0

US-09-712-338-1 (1-1668) x Q9FH06 (1-469)

QY 295 ACATTGTTGAATGGTGGCCCTGA 321

Db 74 ThrLeuTriLeuAsnGlyProGly 82

Search completed: November 21, 2003, 17:54:35

Job time : 83.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:44:24 ; Search time 19 Seconds
(without alignments)
8256.899 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 555
Sequence: 1 atgcgtggtagaattctt.....ccagtgttgtagcatag 1668

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Word size: 1

Total number of hits satisfying chosen parameters: 255550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO spoel.p/US09712338/runat.17112003.170225.1076/app_query.fasta.1.1863
-DB=SwissProt 41 -OPMT=fastan -SUFFIX=oli.rsp -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1.1.25 @runat.17112003.170225.10761 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt 41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	1.8	423	1	CPS1_PENJA
2	9	1.6	508	1	YBY9_YEAST
3	8	1.4	36	1	P3BY_ODOSI
4	8	1.4	292	1	YCL5_PYRO
5	8	1.4	356	1	CKRV_MOUSE
6	8	1.4	423	1	CBP2_WHEAT
7	8	1.4	436	1	CBP2_HORVU
8	8	1.4	470	1	YSS2_CAEEL
9	8	1.4	474	1	PRTP_MOUSE
10	8	1.4	476	1	CBP2_HORVU
11	8	1.4	480	1	PRTP_HUMAN
12	8	1.4	499	1	CBP1_HORVU
13	8	1.4	510	1	CBP1_ORYGA
14	8	1.4	516	1	CBP2_HORVU
15	8	1.4	523	1	CBPY_PICPA
16	8	1.5	526	1	VSM4_TRYEB
17	8	1.4	532	1	CBPY_YEAST
18	8	1.4	542	1	CBPY_CANAL

19	8	1.4	574	1	YPP3_CAEEL
20	8	1.4	605	1	WSC4_YEAST
21	8	1.5	760	1	EZ_DROME
22	8	1.4	1002	1	CBFY_SCHPO
23	8	1.4	2145	1	CYAA_PODAN
24	8	1.4	4349	1	FAT2_HUMAN
25	7	1.3	25	1	YCX8_ODOSI
26	7	1.3	66	1	PSAE_FREDI
27	7	1.3	70	1	PSAE_CYPAA
28	7	1.3	88	1	YPB4_ECOLI
29	7	1.3	91	1	YLS9_CAEEL
30	7	1.3	110	1	RLA4_YEAST
31	7	1.3	112	1	H34_STYLE
32	7	1.3	113	1	H38_STYLE
33	7	1.3	114	1	H31_STYLE
34	7	1.3	114	1	H33_STYLE
35	7	1.3	114	1	H35_STYLE
36	7	1.3	114	1	H36_STYLE
37	7	1.3	114	1	H37_STYLE
38	7	1.3	114	1	H39_STYLE
39	7	1.3	117	1	ARR2_ECOLI
40	7	1.3	117	1	SMD1_SCHPO
41	7	1.3	124	1	PA25_AGKHP
42	7	1.3	126	1	YE66_AQUAE
43	7	1.3	132	1	RS8_MYCLE
44	7	1.3	133	1	YJFI_ECOLI
45	7	1.3	134	1	H3_CHLRE

ALIGNMENTS

RESULT 1					
CPS1_PENJA		STANDARD;	PRT;	423 AA.	
ID	CPS1_PENJA				
AC	P34946;				
DT	01-FEB-1994 (Rel. 28, Created)				
DT	01-FEB-1994 (Rel. 28, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Carboxypeptidase S1 (EC 3.4.16.6)				
OS	Penicillium janthinellum (Penicillium vitale)				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.				
OX	NCBI_TaxID=5079;				
RN	[1]				
RP	SEQUENCE.				
RX	MEDLINE=94039747; PubMed=8224168;				
RA	Svendsen I., Hofmann T., Endrizzi J., Remington S.J., Breddam K.;				
RT	"The primary structure of carboxypeptidase S1 from Penicillium janthinellum";				
RL	FEBS Lett. 333:39-43 (1993).				
CC	-!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine or lysine residue.				
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.				
DR	PIR; S38953; S38953.				
DR	HSSP; P00729; 1YSC.				
DR	MEROPS; S10.008; --				
DR	InterPro; IPR000379; Ser estrs site.				
DR	InterPro; IPR001563; Serine carboxpt.				
DR	Pfam; PF00450; serine carboxpt; 1.				
DR	PRINTS; PR00724; CRBOXYPTASEC.				
DR	ProDom; PD001189; Serine carboxpt; 1.				
DR	PROSITE; PS00131; CARBOXYPEPT SER SER; 1.				
DR	PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.				
DR	Hydrolase; Carboxypeptidase; Glycoprotein.				
FT	ACT SITE 143 143				
FT	ACT SITE 340 340				
FT	ACT SITE 397 397				
FT	BINDING 343 343				
FT	BINDING 398 398				
FT	DISULFID 8 68				
FT	DISULFID 55 300				
FT	DISULFID 223 246				
FT	DISULFID 230 239				

P2716	caenorhabdi
P38739	saccharomyc
P42124	drosophila
Q13849	schizosacch
Q1513	podospora a
Q9nyq8	homo sapien
P23809	odontella s
P48114	cyamophora
P03853	escherichia
P03434	caenorhabdi
P02400	saccharomyc
P81198	stylyonchia
P81202	stylyonchia
P81195	stylyonchia
P81197	stylyonchia
P81199	stylyonchia
P81200	stylyonchia
P81201	stylyonchia
P81196	stylyonchia
P52144	escherichia
O42661	schizosacch
O42189	agkistrodon
O67446	aquifex aeo
O32997	mycobacteri
P39291	escherichia
P50564	chlamydomon

FT CARBOHYD 200 200 N-LINKED (GLCNAC...).
SQ SEQUENCE 423 AA; 46496 MW; 55F3333608B12729 CRC64;
Alignment Scores:
Pred. No.: 0.296 Length: 423
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 1 Gaps: 0
US-09-712-338-1 (1-1668) x CPS1_PENJA (1-423)
QY 679 ACGGAGGCTATGAGGCGACATGTCCT 708
Db 141 ThrGluSerTyrGlyGlyHisTyrGlyPro 150
RESULT 2
YBY9 YEAST
ID YBY9 YEAST STANDARD; PRT; 508 AA.
AC P38109;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine carboxypeptidase in ESRI-IRAI intergenic region
DE (EC 3.4.16.-).
GN YBR139W OR YBR1015.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=9437817; PubMed=8091856;
RA Becam A.-M., Cullin C., Gryzbowska E., Lacroute F., Nasr F.,
RA Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
RA Herbert C.J.;
RT "The sequence of 29.7 kb from the right arm of chromosome II reveals
RT 13 complete open reading frames, of which ten correspond to new
RT genes.";
RL Yeast 10:S1-S11(1994).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=95042830; PubMed=7954890;
RA Nasr F., Becam A.-M., Gryzbowska E., Zagulski M., Slonimski P.P.,
RA Herbert C.J.;
RT "An analysis of the sequence of part of the right arm of chromosome
RT II of S. cerevisiae reveals new genes encoding an amino-acid permease
RT and a carboxypeptidase.";
RL Curr. Genet. 26:1-7(1994).
RN [2]
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75891; CAAS3497.1; -;
CC EMBL; Z36008; CAAB5097.1; -;
CC PIR; S46008; S46008.
CC HSP; P00729; 1CPY.
CC MR0PS; S10.UPW.-;
CC SGD; S0000343; YBR139W.
CC InterPro; IPR000379; Ser_estrs.site.
CC InterPro; IPR001563; Serine_carbpept.
CC Pfam; PF00450; serine_carbpept; 1.
CC PRINTS; PR00724; CR00XYPTASEC.
CC PRODOM; PD001189; Serine_carbpept; 1.
CC PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hypothetical protein; Hydrolase; Carboxypeptidase.
FT ACT_SITE 219 219 BY SIMILARITY.
FT ACT_SITE 415 415 BY SIMILARITY.
FT ACT_SITE 474 474 BY SIMILARITY.
SQ SEQUENCE 508 AA; 57639 MW; AAB2806C8EE2EBE1 CRC64;
Alignment Scores:
Pred. No.: 2.85 Length: 508
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 1 Gaps: 0
US-09-712-338-1 (1-1668) x YBY9_YEAST (1-508)
QY 439 CAGCCATTGGGAGTCGGCTTTTCAT 465
Db 171 GlnProLeuglyValGlyPheSerTyr 179
RESULT 3
PSBY ODOI
ID PSBY ODOI STANDARD; PRT; 36 AA.
AC P49543;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II protein Y.
GN (PSBY-A OR PSBY-1 OR YCF32.1) AND (PSBY-B OR PSBY-2 OR YCF32.2).
OS Odontella sinensis (Marine centric diatom).
OG Chloroplast.
OC Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Bidulphiophycidae; Eupodiscales; Eupodiscaeae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -!- FUNCTION: MANGANESE-BINDING POLYPEPTIDE WITH L-ARGININE
CC METABOLIZING ENZYME ACTIVITY. COMPONENT OF THE CORE OF PHOTOSYSTEM
CC II (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- SIMILARITY: BELONGS TO THE PSBY FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z67753; CAA91612.1; -;
CC EMBL; Z67753; CAA91666.1; -;
CC PIR; S78239; S78239.
CC KW Photosystem II; Chloroplast; Transmembrane; Thylakoid.
CC FT DOMAIN 1 6 POTENTIAL.
CC FT TRANSMEM 7 23 STROMAL (POTENTIAL).
CC FT DOMAIN 24 36 STROMAL (POTENTIAL).
CC SQ SEQUENCE 36 AA; 3921 MW; 22989D3E7DAE2B63 CRC64;
Alignment Scores:
Pred. No.: 45 Length: 36
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0
US-09-712-338-1 (1-1668) x PSBY_ODOI (1-36)

Qy 34 TTGGTTCAGCCAGTGGCCCTT 57
 Db 13 LeuValAlaIleSerTrpAlaLeu 20

RESULT 4

YC15_PVRHO ID YC15_PVRHO STANDARD; PRT; 292 AA.
 AC OS8966;
 DT 13-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical ABC transporter permease protein PH1215.
 GN PH1215 OR PHK039.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Roki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 CC -!- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM PH1214/15/16. PROBABLY RESPONSIBLE FOR THE TRANSLOCATION
 CC OF THE SUBSTRATE ACROSS THE MEMBRANE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM PERMEASE FAMILY. MALFG SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AP000005; BAA30315.1; -;
 DR PIR; A71065; A71065.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp; 1.
 DR PROSITE; PS00402; BPD_TRANSPO_INN_MEMBER; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 261 281 POTENTIAL.
 SQ SEQUENCE 292 AA; 33407 MW; 5B968812C0B99F68 CRC64;

Alignment Scores:
 Pred. No.: 31.1 Length: 292
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 1 Gaps: 0

US-09-712-338-1 (3-1668) x YC15_PVRHO (1-292)

Qy 755 GTAGTGTAAATGGTTCAGCTTA 778

Db 219 ValValLeuMetValPheSerLeu 226

RESULT 5

CKRV_MOUSE ID CKRV_MOUSE STANDARD; PRT; 356 AA.
 AC P51676;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor 1-like protein 1 (Macrophage inflammatory
 DE protein-1 alpha receptor-like 1).
 GN CCR1L1 OR CNKBR1L1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=95340546; PubMed=7542241;
 RA Gao J.-L., Murphy P.M.;
 RT "Cloning and differential tissue-specific expression of three mouse
 RT beta chemokine receptor-like genes, including the gene for a
 RT functional macrophage inflammatory protein-1 alpha receptor.";
 RL J. Biol. Chem. 270:17494-17501(1995).
 CC -!- FUNCTION: PROBABLE RECEPTOR FOR A C-C TYPE CHEMOKINE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: DETECTED IN THE SPLEEN, LIVER AND LEUKOCYTES.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U28405; AAA89154.1; -;
 DR PIR; I49340; I49340.
 DR MGD; MGI:104617; Ccr1l1.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECP F1_2; 1.
 KW G-protein coupled receptor; Transmembrane.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 60 1 (POTENTIAL).
 FT DOMAIN 61 67 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 68 92 2 (POTENTIAL).
 FT DOMAIN 93 108 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 109 130 3 (POTENTIAL).
 FT DOMAIN 131 147 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 148 172 4 (POTENTIAL).
 FT DOMAIN 173 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 224 5 (POTENTIAL).
 FT DOMAIN 225 240 6 (POTENTIAL).
 FT TRANSMEM 241 265 7 (POTENTIAL).
 FT DOMAIN 266 282 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 283 306 7 (POTENTIAL).
 FT DOMAIN 307 356 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 107 184 BY SIMILARITY.
 SQ SEQUENCE 356 AA; 40934 MW; 58C01ABA8D7D4B06 CRC64;

Alignment Scores:
 Pred. No.: 30.1 Length: 356
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 1 Gaps: 0

US-09-712-338-1 (1-1668) x CKRV_MOUSE (1-356)


```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0

US-09-712-338-1 (1-1668) x CBP2_WHEAT (1-423)
QY 298 TTGTGTTGAATGTCGCCCTGGA 321
  |||||
Db 57 LeuTriPLeuAsnGlyGlyProGly 64

RESULT 7
CP22_HORVU STANDARD; PRT; 436 AA.
ID AC P55748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II-2 precursor (EC 3.4.16.6) (CP-MII.2)
DE (Fragment).
GN CXP2-2.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Alexis; TISSUE=Grain;
RX MEDLINE=94336715; PubMed=7520177;
RA Dal Degun F., Rocher A., Cameron-Mills V., von Wettstein D.;
RT "The expression of serine carboxypeptidases during maturation and
  germination of the barley grain.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213 (1994).
CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
  or lysine residue.
CC -!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
  COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. LOW
  LEVELS IN THE DEVELOPING ALEURONE AND EMBRYO. ALSO FOUND IN THE
  ROOTS AND SHOOTS OF THE GROWING SEEDLING.
CC -!- PTM: THE LINKER PEPTIDE IS ENDOPEPTOLYTICALLY EXCISED DURING
  ENZYME MATURATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X78878; CAB59202.1; -.
DR HSSP; P08819; 1WHT.
DR InterPro; IPR000379; Ser_estr_site.
DR InterPro; IPR001563; Serine carboxpept.
DR Pfam; PF00450; serine carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Zymogen; Multigene family.
FT NON TER 1
FT CHAIN <1 256
FT PROPEP 257 270
FT CHAIN 271 436
FT ACT_SITE 149 149
FT ACT_SITE 350 350
FT ACT_SITE 403 403
FT DISULFID 56 313
FT DISULFID 217 229
FT DISULFID 253 281
FT CARBOHYD 107 107

SERINE CARBOXYPEPTIDASE II-2, CHAIN A.
LINKER PEPTIDE (BY SIMILARITY).
SERINE CARBOXYPEPTIDASE II-2, CHAIN B.
BY SIMILARITY.
BY SIMILARITY.
INTERCHAIN (BY SIMILARITY).
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC...) (POTENTIAL).

Alignment Scores:
Pred. No.: 1
Score: 1
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Mismatches: 0
Conservative: 0
Gaps: 0

SQ SEQUENCE 436 AA; 48952 MW; E0F82D97E0C34DC9 CRC64;
Alignment Scores:
Pred. No.: 29
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Mismatches: 0
Conservative: 0
Indels: 0
Gaps: 0

US-09-712-338-1 (1-1668) x CP22_HORVU (1-436)
QY 298 TTGTGTTGAATGTCGCCCTGGA 321
  |||||
Db 48 LeuTriPLeuAsnGlyGlyProGly 55

RESULT 8
YSS2_CAEEL STANDARD; PRT; 470 AA.
ID AC Q09921;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine carboxypeptidase K10B2.2 precursor (EC 3.4.16.-).
GN K10B2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U28730; AAA68259.1; -.
DR PIR; T16606; T16606.
DR HSSP; P10619; 1IIV.
DR MEROPS; S10.002; -.
DR WormPep; K10B2.2; CE02009.
DR InterPro; IPR000379; Ser_estr_site.
DR InterPro; IPR001563; Serine carboxpept.
DR Pfam; PF00450; serine carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 2.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hypothetical protein; Hydrolase; Carboxypeptidase; Glycoprotein;
  Signal.
FT SIGNAL 1 19
FT CHAIN 20 470
FT ACT_SITE 169 169
FT ACT_SITE 380 380
FT ACT_SITE 441 441
FT CARBOHYD 132 132
FT CARBOHYD 316 316
FT CARBOHYD 396 396
FT SEQUENCE 470 AA; 53158 MW; CCC2DACB75EF30FC CRC64;

PUTATIVE SERINE CARBOXYPEPTIDASE K10B2.2.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).

Alignment Scores:
Pred. No.: 28.6
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Mismatches: 0
Conservative: 0
Mismatches: 0
```

Query Match: 1.44% Indels: 0
DB: Caps: 0

US-09-712-338-1 (1-1668) x YSS2_CAEEL (1-470)

QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321
|||||
Db 71 LeuTrpLeuAsnGlyGlyProGly 78

RESULT 9
PRTP MOUSE
ID PRTP MOUSE STANDARD; PRT; 474 AA.
AC P16675; Q8VBF6;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lysosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A)
DE (Carboxypeptidase C) (protective protein for beta-galactosidase).
GN PPG8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90170982; PubMed=2106523;
RA Galjart N.J., Gilleman N., Meijer D., D'Azzo A.;
RT "Mouse 'protective protein'. cDNA cloning, sequence comparison, and
RT expression.";
RL J. Biol. Chem. 265:4678-4684 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Carninci P., Prange C.,
RA Richards S., Worley K.C., Hale S., Garcia A.J., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: PROTECTIVE PROTEIN APPEARS TO BE ESSENTIAL FOR BOTH THE
CC ACTIVITY OF BETA-GALACTOSIDASE AND NEURAMINIDASE, IT ASSOCIATES
CC WITH THESE ENZYMES AND EXERTS A PROTECTIVE FUNCTION NECESSARY FOR
CC THEIR STABILITY AND ACTIVITY. THIS PROTEIN IS ALSO A
CC CARBOXYPEPTIDASE AND CAN DEAMIDATE TACHYKININS.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
CC broad specificity.
CC -!- SUBUNIT: HETERODIMER OF A 32 kDa AND 20 kDa CHAINS HELD TOGETHER
CC BY DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

EMBL; J05261; AAA39982.1; -.
DR EMBL; BC018534; AH18534.1; -.
DR PIR; A35732; A35732.
DR HSP; P10619; LIVY.
DR MEROPS; S10.002; -.
DR MGD; MGI:97748; Ppqb.
DR InterPro; IPR000373; Ser estrs site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carbpept; 2.
DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal; Lysosome.
FT SIGNAL 1 23
FT CHAIN 24 474
FT CHAIN 24 320
FT CHAIN 321 474
FT ACT_SITE 173 173
FT ACT_SITE 394 394
FT ACT_SITE 451 451
FT DISULFID 83 356
FT DISULFID 235 251
FT DISULFID 236 241
FT DISULFID 276 325
FT CARBOHYD 140 140
FT CARBOHYD 327 327
FT CONFLICT 425 425
SQ SEQUENCE 474 AA; 53844 MW; 8F4D5A7F4FEB6C6CE CRC64;

Alignment Scores:
Pred. No.: 28.6 Length: 474
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: Gaps: 0

US-09-712-338-1 (1-1668) x PRTP_MOUSE (1-474)

QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321
|||||
Db 75 LeuTrpLeuAsnGlyGlyProGly 82

RESULT 10
ID CBP2 HORVU STANDARD; PRT; 476 AA.
AC P08618; P93177;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II precursor (EC 3.4.16.6) (Carboxypeptidase
DE D) (CP-MII) [Contains: Serine carboxypeptidase II chain A; Serine
DE carboxypeptidase II chain B].
GN CBP2 OR CBP.2
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Igri. TISSUE=Etisolated leaf;
RA Rocher A., Lok F., Cameiron-Mills V., von Wettstein D.;
RL "The gene family of serine carboxypeptidases in barley";
RL Submitted (NOV-1996) to the EMBL/GenBank/DDJ databases.
RN [2]
RP SEQUENCE OF 35-294 AND 314-472.
RA Soerensen S.B., Svendsen I., Breddam K.;
RA "Primary structure of carboxypeptidase II from malted barley.";
RL Carlsberg Res. Commun. 52:285-295 (1987).
CC -!- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF SMALL PEPTIDES (2-
CC 5 RESIDUES) OR IN THE DEGRADATION OF STORAGE PROTEINS IN THE
CC

CC CC EMBRYO.

CC CC -1- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine

CC CC or lysine residue.

CC CC -1- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS

CC CC COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND.

CC CC -1- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.

CC CC -1- DEVELOPMENTAL STAGE: SIMULTANEOUSLY PRESENT IN ALEURONE AND

CC CC ENDOSPERM BETWEEN 20 AND 30 DAYS POSTANTHESIS. ACCUMULATES IN THE

CC CC DEVELOPING GRAIN AND IS STORED IN ITS ACTIVE FORM IN THE MATURE

CC CC GRAIN. ALSO FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.

CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC CC -----

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC CC the European Bioinformatics Institute. There are no restrictions on its

CC CC use by non-profit institutions as long as its content is in no way

CC CC modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC CC or send an email to license@isb-sib.ch).

CC CC -----

CC CC EMBL: Y09602; CAA70815.1; -

CC CC PIR: T05701; T05701.

CC CC HSSP: P08819; 1WHT.

CC CC DR MEROPS: S10.005; -

CC CC DR InterPro: IPR001563; Serine_carbpept.

CC CC DR Pfam: PF00450; serine_carbpept; 1.

CC CC DR PRINTS: PR00724; CRBOXYPTASEC.

CC CC DR PROSITE: PD001189; Serine_carbpept; 1.

CC CC DR PROSITE: PS00131; CARBOXYPEPT_SER_SER; 1.

CC CC DR PROSITE: PS00560; CARBOXYPEPT_SER_HIS; 1.

CC CC KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal;

CC CC KW Multigene family.

CC CC FT SIGNAL 1 34

CC CC FT CHAIN 35 294

CC CC FT PROPEP 295 313

CC CC FT CHAIN 314 476

CC CC FT ACT_SITE 190 190

CC CC FT ACT_SITE 390 390

CC CC FT ACT_SITE 443 443

CC CC FT DISULFID 97 333

CC CC FT DISULFID 254 266

CC CC FT DISULFID 290 320

CC CC FT CARBOHYD 148 148

CC CC FT CARBOHYD 159 159

CC CC FT CARBOHYD 291 291

CC CC FT CARBOHYD 341 341

CC CC FT CARBOHYD 347 347

CC CC FT CARBOHYD 352 352

CC CC FT CARBOHYD 472 472

CC CC FT MOD_RES 314 314

CC CC FT VARIANT 351 352

CC CC FT CONFLICT 181 181

CC CC SQ SEQUENCE 476 AA; 52625 MW; 18685725B1A6B5E4 CRC64;

CC CC -----

CC CC Alignment Scores:

CC CC Pred. No.: 28.6 Length: 476

CC CC Score: 8.00 Matches: 8

CC CC Percent Similarity: 100.00% Conservative: 0

CC CC Best Local Similarity: 100.00% Mismatches: 0

CC CC Query Match: 1.44% Indels: 0

CC CC DB: 1 Gaps: 0

CC CC -----

CC CC US-09-712-338-1 (1-1668) x CBP2_HORVU (1-476)

CC CC Qy 298 TTGTGGTGAATGGTGGCCCTGGA 321

CC CC Db 89 LeuTrpLeuAsnGlyGlyProGly 96

CC CC =====

CC CC RESULT 11

CC CC PRTP_HUMAN

CC CC ID PRTP_HUMAN STANDARD; PRT; 480 AA.

CC CC AC P10619; Q96KJ2; Q9BW68;

CC CC DT 01-JUL-1989 (Rel. 11, Created)

DT DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE DE Lysosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A)

DE DE (Carboxypeptidase C) (Protective protein for beta-galactosidase).

GN GN PCPB.

OS OS Homo sapiens (Human).

OC OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

OC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX OX NCBI_TaxID=9606;

RN RN [1]

RP RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX RX MEDLINE=88311078; PubMed=3136930;

RA RA Galjart N.J., Gillemans N., Harris A., van de Horst G.T.J.,

RA RA Verheijen F.W., Galjaard H., D'Azio A.,

RT RT "Expression of cDNA encoding the human 'protective protein'

RT RT associated with lysosomal beta-galactosidase and neuraminidase:

RT RT homology to yeast proteases.";

RL RL Cell 54:755-764(1988).

RN RN [2]

RP RP SEQUENCE FROM N.A.

RX RX MEDLINE=21638749; PubMed=11780052;

RA RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,

RA RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,

RA RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA RA Coulson A., Coville G.J., Collier R.E., Connor R.E., Corby N.R.,

RA RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA RA Hammond D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

RA RA Levaeslatho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,

RA RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RA RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,

RA RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,

RA RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA RA Tracey A., Tomans A.C., Vaudin M., Wall M., Wallis J.M.,

RA RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,

RA RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RT RT Rogers J.;

RL RL "The DNA sequence and comparative analysis of human chromosome 20.";

RL RL Nature 414:865-871(2001).

RN RN [3]

RP RP SEQUENCE FROM N.A.

RX RX TISSUE=Skin;

RA RA MEDLINE=22388257; PubMed=12477932;

RA RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA RA Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Sanchez A.,

RA RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,

RA RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,

RA RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,

RT RT "Generation and initial analysis of more than 15,000 full-length

RT RT human and mouse cDNA sequences.";

RL RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN RN [4]

RP SEQUENCE OF 29-53 AND 327-351.
RC TISSUE-Plasminogen activator.
RX MEDLINE=90293074; PubMed=1694176;
RA Jackman H.L., Tan F., Tamei H., Beurling-Harbury C., Li X.-Y.,
RA Skidgel R.A., Erdos E.G.;
RT "A peptidase in human platelets that deamidates tachykinins. Probable
RT identity with the lysosomal 'protective protein'.";
RL J. Biol. Chem. 265:11265-11272(1990).
RN [5]
RN FUNCTION, AND MUTAGENESIS.
RP MEDLINE=91317848; PubMed=1907282;
RX Galjart N.J., Morreau H., Willemsen R., Gillemans N., Bonten E.J.,
RA D'Azzo A.;
RT "Human lysosomal protective protein has cathepsin A-like activity
RT distinct from its protective function.";
RL J. Biol. Chem. 266:14754-14762(1991).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RP MEDLINE=96164441; PubMed=8591035;
RX Rudenko G., Bonten E., D'Azzo A., Hol W.G.J.;
RA "Three-dimensional structure of the human 'protective protein':
RT structure of the precursor form suggests a complex activation
RT mechanism.";
RL Structure 3:1249-1259(1995).
RN [7]
RN VARIANT GALACTOSIALIDOSIS VAL-440.
RP MEDLINE=92097522; PubMed=1756715;
RX Zhou X.-Y., Galjart N.J., Willemsen R., Gillemans N., Galjaard H.,
RA D'Azzo A.;
RT "A mutation in a mild form of galactosialidosis impairs dimerization
RT of the protective protein and renders it unstable.";
RL EMBO J. 10:4041-4048(1991).
RN [8]
RN VARIANTS GALACTOSIALIDOSIS ARG-49; ARG-65; LEU-90; ASN-249 AND
RP CYS-395.
RX MEDLINE=93293970; PubMed=8514852;
RA Shimamoto M., Fukuhara Y., Itoh K., Oshima A., Sakuraba H., Suzuki Y.;
RT "Protective protein gene mutations in galactosialidosis.";
RL J. Clin. Invest. 91:2393-2398(1993).
CC -!- FUNCTION: PROTECTIVE PROTEIN APPEARS TO BE ESSENTIAL FOR BOTH THE
CC ACTIVITY OF BETA-GALACTOSIDASE AND NEURAMINIDASE. IT ASSOCIATES
CC WITH THESE ENZYMES AND EXERTS A PROTECTIVE FUNCTION NECESSARY FOR
CC THEIR STABILITY AND ACTIVITY. THIS PROTEIN IS ALSO A
CC CARBOXYPEPTIDASE AND CAN DEAMIDATE TACHYKININS.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
CC broad specificity.
CC -!- SUBUNIT: HETERODIMER OF A 32 kDa AND 20 kDa CHAINS HELD TOGETHER
CC BY DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- DISEASE: Defects in PPGB are the cause of galactosialidosis
CC [MIM:256540], an autosomal recessive disease.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M22960; AAA36476.1; -;
CC EMBL; AL008726; CAA15501.1; -;
CC EMBL; BC000597; AAH00597.1; -;
CC PIR; A31589; A31589;
CC PDB; 1IVY; 2I-APR-97.
CC MEROPS; S10.002; -;
CC Genew; HGNC:9251; PPGB.
CC MIM; 256540; -;
CC GO; GO:0005783; C:endoplasmic reticulum; TAS.
CC GO; GO:0004180; F:carboxypeptidase activity; TAS.
CC GO; GO:0008047; F:enzyme activator activity; TAS.
CC GO; GO:0005478; F:intracellular transporter activity; TAS.

GO; GO:0006886; P:intracellular protein transport; TAS.
DR InterPro; IPR000379; Ser_estr site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine_carbpept; 2.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR KX Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal; Lysosome;
KW 3D-structure; Disease mutation.
FT SIGNAL 1 28
FT CHAIN 29 480 LYSOSOMAL PROTECTIVE PROTEIN.
FT CHAIN 29 326 32 kDa CHAIN.
FT CHAIN 327 480 20 kDa CHAIN.
FT ACT SITE 178 178
FT ACT SITE 400 400
FT ACT SITE 457 457
FT DISULFID 88 362
FT DISULFID 240 256
FT DISULFID 241 246
FT DISULFID 281 331
FT CARBOHYD 145 145
FT CARBOHYD 333 333
FT VARIANT 49 49
FT VARIANT 65 65
FT VARIANT 90 90
FT VARIANT 249 249
FT VARIANT 395 395
FT VARIANT 440 440
FT MUTAGEN 178 178
FT MUTAGEN 457 457
FT CONFLICT 19 19
FT CONFLICT 56 56
FT HELIX 31 33
FT STRAND 34 34
FT STRAND 39 39
FT STRAND 49 55
FT TURN 58 59
FT STRAND 60 67
FT TURN 73 75
FT STRAND 78 82
FT TURN 85 87
FT STRAND 88 88
Alignment Scores:
Pred. No.: 28.5 Length: 480
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0
US-09-712-338-1 (1-1668) x PRTP_HUMAN (1-480)
QY 298 TTGTGTTGTAATGTCGCGCTGGA 321
DB 80 LeutripLeuAsnGlyGlyProGly 87
RESULT 12
ID CBP1_HORVU STANDARD; PRT; 499 AA.
AC P07519; P07520;
DT 01-APR-1998 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase
 DE C) (CP-MI).
 GN CBPI OR CXP;1.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 CX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aleurone;
 RA Rocher A., Lok F., Cameron-Mills V., von Wettstein D.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 88-499 FROM N.A.
 RX MEDLINE=88296749; PubMed=3403516;
 RA Doan N.P., Fincher G.B.;
 RT "The A- and B-chains of carboxypeptidase I from germinated barley
 RT originate from a single precursor polypeptide.";
 RL J. Biol. Chem. 263:11106-11110(1988).
 RN [3]
 RP SEQUENCE OF 31-296 AND 352-499.
 RA Sorensen S.B., Breddam K., Svendsen I.;
 RT "Primary structure of carboxypeptidase I from malted barley.";
 RL Carlsberg Res. Commun. 51:475-485(1986).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF SMALL PEPTIDES (2-
 CC 5 RESIDUES) OR IN THE DEGRADATION OF STORAGE PROTEINS IN THE
 CC EMBRYO.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 CC broad specificity.
 CC -1- SUBUNIT: CARBOXYPEPTIDASE I IS A DIMER, WHERE EACH MONOMER IS
 CC COMPOSED OF TWO CHAINS LINKED BY DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.
 CC -1- DEVELOPMENTAL STAGE: AFTER ONE DAY OF GERMINATION, MAINLY FOUND IN
 CC THE SCUTELLUM OF THE DEVELOPING GRAIN; BARELY DETECTABLE AFTER
 CC FOUR DAYS, AND ABSENT FROM THE MATURE GRAIN. A LOWER LEVEL OF
 CC EXPRESSION IS SEEN IN THE ALEURONE BOTH DURING DEVELOPMENT AND
 CC GERMINATION.
 CC -1- PTM: THREE DISULFIDE BONDS ARE PRESENT.
 CC -1- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING
 CC ENZYME MATURATION.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Y09603; CAA70816.1; -;
 DR EMBL; J03897; AAA32940.1; -;
 DR PIR; T05367; CPBS.
 DR HSSP; P08819; 1WHT.
 DR MEROPS; S10.004; -;
 DR InterPro; IPR000379; Ser esters site.
 DR InterPro; IPR001563; Serine carboxpept.
 DR Pfam; PF00450; serine carboxpept; 1.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine carboxpept; 2.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 DR Hydrolase; Carboxypeptidase; Zymogen; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 296
 FT PROPEP 297 351
 FT CHAIN 352 499
 FT ACT_SITE 188 188
 FT ACT_SITE 423 423
 FT ACT_SITE 476 476
 FT CARBOHYD 148 148
 FT CARBOHYD 262 262
 FT N-LINKED (GLCNAC. . .).
 FT BY SIMILARITY.

FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .).
 FT SITE 497 499 MICROBODY TARGETING SIGNAL (POTENTIAL).
 FT CONFLICT 102 102 H -> P (IN REF. 3).
 SQ SEQUENCE 499 AA; 54096 MW; 9C6674B14D9DB9BF CRC64;
 Alignment Scores:
 Pred. No.: 28.3 Length: 499
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 1 Gaps: 0
 US-09-712-338-1 (1-1668) x CBPI_HORVU (1-499)
 QY 298 TTGTGTTGAATGGTGCCTGGA 321
 Db 84 LeuTriLeuAsnGlyGlyProGly 91
 RESULT 13
 CBPI_ORYSA
 ID CBPI_ORYSA STANDARD; PRT; 510 AA.
 AC P37890;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase
 DE C).
 GN CBPI.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzaceae; Oryza.
 CX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Yukihikari;
 RX MEDLINE=94213891; PubMed=8161571;
 RA Washio K., Ishikawa K.;
 RT "Cloning and sequencing of the gene for type I carboxypeptidase in
 RT rice.";
 RL Biochim. Biophys. Acta 1199:311-314(1994).
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 CC broad specificity.
 CC -1- PTM: THREE DISULFIDE BONDS ARE PRESENT (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D17586; BAA04510.1; -;
 DR PIR; S43516; S43516.
 DR HSSP; P08819; 1WHT.
 DR MEROPS; S10.004; -;
 DR Gramene; P37890; -;
 DR InterPro; IPR000379; Ser esters site.
 DR InterPro; IPR001563; Serine carboxpept.
 DR Pfam; PF00450; serine carboxpept; 1.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine carboxpept; 2.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 DR Hydrolase; Carboxypeptidase; Zymogen; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 36
 FT PROPEP 37 510
 FT CHAIN 194 194
 FT ACT_SITE 434 434
 FT ACT_SITE 434 434
 FT BY SIMILARITY.

FT ACT SITE 487 487 BY SIMILARITY.
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 508 510 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 510 AA; 55709 MW; 28896247FA1371CF CRC64;

Alignment Scores: 28.2 Length: 510
Pred. No.: 8.00 Matches: 8
Score: 8.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.44% Gaps: 0
DB: 1

US-09-712-338-1 (1-1668) x CBPI_ORYSA (1-510)
QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321
DB 90 LeuTrpLeuAenGlyGlyProGly 97

RESULT 14
CP23 HORVU
ID CP23 HORVU STANDARD; PRT; 516 AA.
AC P52711;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II-3 precursor (EC 3.4.16.6) (CP-MII.3).
GN EXP.2-3.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Alexis; TISSUE=Grain;
RX MEDLINE=94336715; PubMed=7520177;
RA Dal Degan F., Rother A., Cameron-Mills V., von Wettstein D.;
RT "The expression of serine carboxypeptidases during maturation and
germination of the barley grain."
RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213(1994).
CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
or lysine residue.
CC -!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. ALSO
FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.
CC -!- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING
ENZYME MATURATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X78877; CAA55478.1; --
PIR; S44191; S44191.
DR HSP; P08819; 1MHT.
DR MEROPS; S10.005; --
DR InterPro; IPR000379; Ser estrs site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXVPTASEC.
DR PRODOM; PD001189; Serine carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal;
KW

KW Multigene family. 20
FT SIGNAL 1 1 POTENTIAL.
FT PROPEP 21 77 POTENTIAL.
FT CHAIN 78 341 SERINE CARBOXYPEPTIDASE II-3, CHAIN A.
FT PROPEP 342 352 LINKER PEPTIDE (BY SIMILARITY). CHAIN B.
FT CHAIN 353 516 SERINE CARBOXYPEPTIDASE II-3, CHAIN B.
FT ACT SITE 236 236 BY SIMILARITY.
FT ACT_SITE 427 427 BY SIMILARITY.
FT ACT_SITE 484 484 BY SIMILARITY.
FT DISULFID 143 391 INTERCHAIN (BY SIMILARITY).
FT DISULFID 300 315 BY SIMILARITY.
FT DISULFID 339 359 BY SIMILARITY.
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 516 AA; 55913 MW; D41AA1C56CF8D355 CRC64;

Alignment Scores: 28.2 Length: 516
Pred. No.: 8.00 Matches: 8
Score: 8.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.44% Gaps: 0
DB: 1

US-09-712-338-1 (1-1668) x CP23_HORVU (1-516)
QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321
DB 135 LeuTrpLeuAenGlyGlyProGly 142

RESULT 15
CBPY PICPA
ID CBPY PICPA STANDARD; PRT; 523 AA.
AC P52710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCV).
GN PRC1.
OS Pichia pastoris (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4922;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 108-142.
RC STRAIN=GTS115;
RX MEDLINE=96381245; PubMed=8789258;
RA Ohi H., Ohtani W., Okazaki N., Furuhashi T.,
RT "Cloning and characterization of the Pichia pastoris PRC1 gene
encoding carboxypeptidase Y."
RL Yeast 12:31-40(1996).
CC -!- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X87987; CAA61240.1; --
PIR; S61713; S61713.
DR HSP; P00729; 1YSC.
DR MEROPS; S10.001; --
DR InterPro; IPR000379; Ser estrs site.
DR InterPro; IPR001563; Serine_carbpept.
KW

DR Pfam; PF00450; serine carboxypept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxypept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 20
FT PROPEP 21 107
FT CHAIN 108 523
FT ACT_SITE 249 249
FT ACT_SITE 441 441
FT ACT_SITE 498 498
FT DISULFID 162 401
FT DISULFID 296 310
FT DISULFID 320 343
FT DISULFID 327 336
FT DISULFID 365 371
FT CARBOHYD 193 193
FT CARBOHYD 271 271
FT CARBOHYD 484 484
FT CARBOHYD 487 487
SQ SEQUENCE 523 AA; 59447 MW; CA6093BDE2E9D178 CRC64;
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).

Alignment Scores:

Pred. No.:	28.1	Length:	523
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.44%	Indels:	0
DB:	1	Gaps:	0

US-09-712-338-1 (1-1668) x CBPY_PICPA (1-523)

QY 298 TTGTGTTGAATGGTGGCCCTGA 321

Db 154 LeuTrpLeuAsnGlyGlyProGly 161

Search completed: November 21, 2003, 17:51:51
Job time : 25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:47:05 ; Search time 31.5 Seconds
(without alignments)
10184.721 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 555
Sequence: 1 atcgctggtagcaattctt.....ccagtggtgtaggcataag 1668

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565704

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2.1/USPTO spool.p/US09712338/runat 17112003 170226 10786/app_query.fasta_1.1863
-DB=PIR 76 -OPMT=fastan -SUFFIX=oli.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1.1.59 @runat 17112003 170226 10786 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR 76.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	1.8	423	2 S38953	carboxypeptidase D
2	10	1.8	460	1 S31516	serine-type carbox
3	9	1.6	405	2 S70150	tniQ protein - Xan
4	9	1.6	405	2 T08521	tniQ protein homol
5	9	1.6	468	2 D84503	probable serine ca
6	9	1.6	479	2 G96501	probable serine ca
7	9	1.6	495	2 S32179	tniQ protein homol
8	9	1.6	508	1 S46008	probable carboxype
9	2	1.6	2105	2 T18968	probable serine-ty
10	8	1.4	36	2 S78239	ycf32 protein - Od
11	8	1.4	91	2 G20064	hypothetical prote
12	8	1.5	116	2 A88450	protein F21H11.1
13	8	1.4	166	2 E71040	hypothetical prote
14	8	1.5	184	2 H69106	conserved hypothet

15	8	1.4	210	2 C84404	hypothetical prote
16	8	1.4	284	2 T49055	serine carboxypept
17	8	1.4	270	2 B75259	hypothetical prote
18	8	1.4	281	2 G82201	hypothetical prote
19	8	1.5	290	2 AD0137	succinate-CoA liga
20	8	1.4	292	2 A71065	probable permease
21	8	1.5	327	2 G83058	probable sugar-bin
22	8	1.4	334	2 T19127	hypothetical prote
23	8	1.4	356	2 T49340	MIP-1 alpha recept
24	8	1.4	423	1 A29639	carboxypeptidase D
25	8	1.4	425	2 F85360	SERINE CARBOXYPEPT
26	8	1.4	425	2 E84631	probable serine ca
27	8	1.4	447	2 G84772	probable serine ca
28	8	1.4	452	2 H84772	probable serine ca
29	8	1.4	456	2 H86406	probable serine ca
30	8	1.4	456	2 A85139	probable serine ca
31	8	1.4	458	2 F84746	hypothetical prote
32	8	1.5	459	2 H97572	probable serine ca
33	8	1.4	465	2 G82444	hypothetical prote
34	8	1.4	465	2 B85358	SERINE CARBOXYPEPT
35	8	1.4	470	2 T16606	probable serine ca
36	8	1.4	470	2 B96637	hypothetical prote
37	8	1.4	474	2 A35732	protective protein
38	8	1.4	476	1 T05701	carboxypeptidase D
39	8	1.4	477	2 D70546	hypothetical prote
40	8	1.4	480	2 A31589	carboxypeptidase C
41	8	1.4	480	2 T50511	serine-type carbox
42	8	1.5	491	2 AI2793	conserved hypothet
43	8	1.4	499	1 CP8HS	carboxypeptidase C
44	8	1.4	501	2 T43081	serine-type carbox
45	8	1.4	502	2 T43180	serin carboxypepti

ALIGNMENTS

RESULT 1

S38953
carboxypeptidase D (EC 3.4.16.6) - Penicillium janthineum
N/A: carboxypeptidase S1
C: Species: Penicillium janthineum
C: Date: 18-Feb-1994 #sequence_revision 19-Apr-1996 #text_change 31-Mar-2000
C: Accession: S38953
R: Svendsen, I.; Hofmann, T.; Endrizzi, J.; Remington, S.J.; Breddam, K.
R: Lett, 333, 39-43, 1993
A: Title: The primary structure of carboxypeptidase S1 from Penicillium janthineum.
A: Reference number: S38953; MUID: 94039747; PMID: 8224168
A: Accession: S38953
A: Status: preliminary
A: Molecule type: protein
A: Residues: 1-423 <SVS>
C: Superfamily: serine carboxypeptidase
C: Keywords: hydrolase; serine carboxypeptidase

Alignment Scores:

Pred. No.: 0.356 Length: 423
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x S38953 (1-423)

Oy 679 ACGAGAGCTATGAGGGGCACTATGGTCCT 708
|||||
Db 141 ThrGluSerTyrGlyGlyHisTyrGlyPro 150
|||||

RESULT 2

S31516
serine-type carboxypeptidase (EC 3.4.16.-) Z precursor - Absidia zychae
C: Species: Absidia zychae
C: Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Jun-2000
C: Accession: S51516; S78013; S78014

Mon Nov 24 13:41:01 2003

us-09-712-338-1.oli.rpr

R;Lee, B.R.; Takeuchi, M.; Kobayashi, Y.
 Curr. Genet. 27, 159-165, 1995
 A;Title: Molecular cloning and sequence analysis of the scpZ gene encoding the serine carboxypeptidase from *Xanthomonas* sp.
 A;Reference number: S51516; MUID:95308538; PMID:7783719
 A;Accession: S51516
 A;Molecule type: DNA
 A;Residues: 1-460 <LEE>
 A;Cross-references: EMBL:D16519; NID:9556466; PIDN:BA03966.1; PID:9995456
 A;Experimental source: strain NRIC 1199
 A;Accession: S78013
 A;Molecule type: protein
 A;Residues: 52-62;90-99;367-381 <LEB>
 A;Accession: S78014
 A;Molecule type: mRNA
 A;Residues: 18-460 <LES>
 C;Genetics:
 A;Gene: scpZ
 A;Introns: 64/3; 95/1; 107/2; 138/3; 234/2; 246/1; 270/1; 343/3; 353/2; 384/3; 437/3
 C;Superfamily: serine carboxypeptidase
 C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase
 F;1-17/Domain: signal sequence #status predicted <SIG>
 F;18-51/Domain: propeptide #status predicted <PR>
 F;52-460/Product: serine-type carboxypeptidase #status experimental <MAT>
 F;40,46,119/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;186/Active site: Ser #status predicted

Alignment Scores:
 Pred. No.: 0.351 Length: 460
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.80% Indels: 0
 DB: 1 Gaps: 0

US-09-712-338-1 (1-1668) x S51516 (1-460)

QY 1456 CCGCTTATGAGCGAGGCCATGAAGTCCCA 1485

Db 430 ArgValTyGluAlaGlyHisGluValPro 439

RESULT 3

S70150
 tniQ protein - *Xanthomonas* sp.
 C;Species: *Xanthomonas* sp.
 C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 11-Jan-2000
 C;Accession: S70150
 R;Kholodii, G.Y.; Mindlin, S.Z.; Bass, I.A.; Yurieva, O.V.; Minakhina, S.V.; Nikiforov, M.; Microbiol. 17, 1189-1200, 1995
 A;Title: Four genes, two ends, and a res region are involved in transposition of Tn5053:
 A;Reference number: S70140; MUID:96130850; PMID:8594337
 A;Accession: S70150
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-405 <KHO>
 A;Cross-references: EMBL:U40585; NID:9710572; PIDN:AA98331.1; PID:9710584
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
 C;Genetics:
 A;Gene: tniQ
 A;Start codon: GTG
 C;Superfamily: *Xanthomonas* tniQ protein
 C;Keywords: transposition

Alignment Scores:
 Pred. No.: 3.86 Length: 405
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x S70150 (1-405)

QY 262 GAAGCCAGACATACCCAGAACTGCA 288

Db 341 GluAlaArgHisAsnProGluThrAla 349
 RESULT 4
 T08521
 tniQ protein homolog - *Enterobacter aerogenes* plasmid R751
 C;Species: *Enterobacter aerogenes*
 C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C;Accession: T08521
 R;Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
 Plasmid 36, 95-111, 1996
 A;Title: Conservation of the genetic switch between replication and transfer genes of *Enterobacter aerogenes* plasmid R751
 A;Reference number: Z16434; MUID:97118926; PMID:8954881
 A;Accession: T08521
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-405 <THO>
 A;Cross-references: EMBL:U67194; NID:91572520; PIDN:AAC64465.1; PID:91572570
 C;Genetics:
 A;Genome: plasmid R751
 C;Superfamily: *Xanthomonas* tniQ protein
 C;Keywords: transposition

Alignment Scores:
 Pred. No.: 3.86 Length: 405
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x T08521 (1-405)

QY 262 GAAGCCAGACATACCCAGAACTGCA 288

Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 5

D84503
 probable serine carboxypeptidase II [imported] - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C;Accession: D84503
 R;Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Kuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: D84503
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-468 <STO>
 A;Cross-references: GB:AE002093; NID:94733981; PIDN:AAD28662.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g12480
 A;Map position: 2
 C;Superfamily: serine carboxypeptidase

Alignment Scores:
 Pred. No.: 3.77 Length: 468
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x D84503 (1-468)

QY 295 ACATTGTGTTGATGTGCGCCCTGCA 321

Db 77 ThrLeuTrpLeuAsnGlyGlyProGly 85

RESULT 6

G96501
 Probable serine carboxypeptidases [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: G96501
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G96501
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-479 <STO>
 A:Cross-references: GB:AE005173; NID:g7523661; PIDN:AAF63101.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F28H19.5
 A:Map position: 1
 C:Superfamily: serine carboxypeptidase

Alignment Scores:

Pred. No.:	3.76	Length:	479
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	2	Gaps:	0

US-09-712-338-1 (1-1668) x G96501 (1-479)

QY 295 ACATGCTGTTGAATGGTGGCCCTGA 321

Db 83 ThrLeuTrpLeuAsnGlyGlyProGly 91

RESULT 7

S32179
 tniQ protein homolog - Klebsiella pneumoniae
 C:Species: Klebsiella pneumoniae
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 C:Accession: S32179
 R:Radstroem, P.; Sundstrom, L.; Swedberg, G.; Flensburg, J.; Skold, O.
 submitted to the EMBL Data Library, March 1993
 A:Description: The integrons of R751 and Tn21 are transposons related to the retroelement
 A:Reference number: S32177
 A:Accession: S32179
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-495 <RAD>
 A:Cross-references: EMBL:X72585; NID:g288626; PIDN:CAA51177.1; PID:g288629
 C:Superfamily: Xanthomonas tniQ protein
 C:Keywords: transposition

Alignment Scores:

Pred. No.:	3.74	Length:	495
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	2	Gaps:	0

US-09-712-338-1 (1-1668) x S32179 (1-495)

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288

Db 431 GiualaArgHisAnProGluThrAla 439

RESULT 8

S46008
 Probable carboxypeptidase C (EC 3.4.16.5) YBR139w - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR1015; hypothetical protein YBR139w
 C:Species: Saccharomyces cerevisiae
 C:Date: 26-Aug-1994 #sequence_revision 05-Jan-1996 #text_change 19-Apr-2002
 C:Accession: S46008; S46581
 R:Becan, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45995
 A:Accession: S46008
 A:Molecule type: DNA
 A:Residues: 1-508 <BEC>
 A:Cross-references: EMBL:Z36008; NID:g536435; PIDN:CAA85097.1; PID:g536436; GSPDB:GN0000
 R:Becan, A.M.; Cullin, C.; Grzybowska, E.; Lacroute, F.; Nasr, F.; Ozier-Kalogeropoulos,
 yeast 10(Suppl.A), S1-S11, 1994
 A:Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 complete
 A:Reference number: S46569; MUID:94378717; PMID:8091856
 A:Accession: S46581
 A:Molecule type: DNA
 A:Residues: 1-508 <BE2>
 A:Cross-references: EMBL:X75891; NID:g456856; PIDN:CAA53497.1; PID:g456869
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0000343
 A:Map position: 2R
 A:Note: MIPS:YBR139w
 C:Superfamily: serine carboxypeptidase
 C:Keywords: hydrolase; serine carboxypeptidase; transmembrane protein
 F:6-22/Domain: transmembrane #status predicted <TMM>
 F:219,415,474/Active site: Ser, Asp, His #status predicted

Alignment Scores:

Pred. No.:	3.73	Length:	508
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	1	Gaps:	0

US-09-712-338-1 (1-1668) x S46008 (1-508)

QY 439 CAGCCATTGGAGTCGGCTTTTCATAT 465

Db 171 GlnProLeuGlyValGlyPheSerTyr 179

RESULT 9

Ti8968
 probable serine-type carboxypeptidase (EC 3.4.16.-) Y16B4A.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
 C:Accession: Ti8968; T23145; T26477
 R:Thomas, K.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19053
 A:Accession: Ti8968
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-2105 <WIL>
 A:Cross-references: EMBL:Z70203; PIDN:CAA94110.1; GSPDB:GN00028; CESP:Y16B4A.2
 R:Lloyd, C.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19697
 A:Accession: T23145
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-2105 <WI2>
 A:Cross-references: EMBL:Z93339; PIDN:CAB07544.1; GSPDB:GN00028; CESP:Y16B4A.2
 R:Wallis, J.
 submitted to the EMBL Data Library, June 1998

A:Reference number: Z20220
A:Accession: T26477
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2105 <WI3>
A:Cross-references: EMBL:AL023825; PIDN:CAA19443.1; GSPDB:GN00028; CESP:Y16B4A.2
A:Experimental source: clone Y16B4A
C:Comment: This protein has 4 repeats of sequence homologous to serine carboxypeptidase.
C:Genetics:
A:Gene: CESP:Y16B4A.2
A:Map position: X
A:Introns: 43/3; 67/2; 179/1; 208/3; 239/2; 281/2; 296/3; 346/2; 444/3; 503/3; 588/3; 681/64/3; 1752/3; 1834/3; 1882/2; 1905/2; 1942/3; 1982/3
C:Keywords: duplication; hydrolase; serine carboxypeptidase

Alignment Scores:
Pred. No.: 3 Length: 2105
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x T18968 (1-2105)
QY 298 TTGTGTTGAATGGTGGCCCTGGAAGC 324
|||||
Db 549 LeuTrpLeuAnGlyGlyProGlySer 557

RESULT 10
S78239
ycf32 protein - Odontella sinensis chloroplast
C:Species: chloroplast Odontella sinensis
C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000
C:Accession: S78239; S78293
R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Frieler, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis
A:Reference number: S78238
A:Accession: S78239
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-36 <KOW>
A:Cross-references: EMBL:Z67753; NID:gl185127; PIDN:CAA91612.1; PID:gl185129
A:Genetics: 5GEN
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
A:Note: 5' gene
A:Accession: S78293
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-36 <KOF>
A:Cross-references: EMBL:Z67753; NID:gl185127; PIDN:CAA91666.1; PID:gl185183
A:Genetics: 5GEN
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
A:Note: 3' gene (c)
A:Genetics: <GEN1>
A:Gene: ycf32
A:Genome: chloroplast
A:Note: gene located on inverted repeat IRA
C:Genetics: <GEN2>
A:Gene: ycf32
A:Genome: chloroplast
A:Note: gene located on inverted repeat IRB
C:Superfamily: hypothetical protein ycf32
C:Keywords: chloroplast

Alignment Scores:
Pred. No.: 60.1 Length: 36
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x S78239 (1-36)
QY 34 TTGTTGACGACCGAGTTGGCCCTT 57
|||||
Db 13 LeuValAlaAlaSerTrpAlaLeu 20

RESULT 11
G90064
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G90064
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogiu, A.; Mizutani-Di, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G90064
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <KUR>
A:Cross-references: GB:BA000018; PID:gl3702540; PIDN:BA843681.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2376

Alignment Scores:
Pred. No.: 52.2 Length: 91
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x G90064 (1-91)
QY 767 GGTTCAGCTTAATTTCAACTCTC 790
|||||
Db 69 ValPheSerLeuIleSerThrLeu 76

RESULT 12
A88450
protein F21H11.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88450
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes.
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/Celegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999.
A:Accession: A88450
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <STO>
A:Cross-references: GB:chr_III; PIDN:AAA19432.1; PID:9506794; GSPDB:GN00021; CESP:F21H11.1
A:Note: coded for by C. elegans cDNA yk13h5.3
C:Genetics:
A:Gene: F21H11.1
A:Map position: 3

Alignment Scores:
Pred. No.: 50.3 Length: 116
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x A88450 (1-116)

QY 1151 GCGTCATGACAGAGTCTTGCC 1128
|||||
Db 99 AlaSerMetThrGluSerPheAla 106
RESULT 13
E71040
hypothetical protein PH1613 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C;Accession: E71040
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: E71040
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-166 <XAW>
A;Cross-references: GB:AP000006; NID:G3236133; PIDN:BAA30725.1; PID:G3258042
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1613

Alignment Scores:
Pred. No.: 47.6 Length: 166
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x E71040 (1-166)

QY 1393 CCCTGAAAGTCAACGGCTGCGAG 1416
|||||
Db 43 ProLeuIysValAsnGlyValGlu 50
RESULT 14
H69106
conserved hypothetical protein MTH1796 - Methanobacterium thermoautotrophicum (strain De
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: H69106
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qi, D.; Spadafora, R.; Vicaires, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: H69106
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-184 <MTH>
A;Cross-references: GB:AE000934; GB:AE000656; NID:G2622924; PIDN:AAB86262.1; PID:G262292
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1796

Alignment Scores:
Pred. No.: 46.8 Length: 184
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x H69106 (1-184)

QY 768 ACCATTACACTACGTTGGCAAT 745
|||||

Db 25 ThrIleAsnThrThrValGlyAsn 32

RESULT 15

C84404

hypothetical protein Vng2543c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: C84404
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, F.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84404
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-210 <STO>
A;Cross-references: GB:AE004437; NID:G10581934; PIDN:AAG20599.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG2543C
C;Superfamily: Pyrococcus horikoshii hypothetical protein PH0010

Alignment Scores:

Pred. No.: 45.9 Length: 210
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x C84404 (1-210)

QY 162 ACCCGGGCAGAGGGCTGCGA 185
|||||

Db 57 ThrArgGlyArgGlyArgLeuArg 64
|||||

Search completed: November 21, 2003, 17:55:48
Job time : 37.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:41:14 ; Search time 52 Seconds
(without alignments)
10182.920 Million cell updates/sec

Title: US-09-712-338-1

Perfect score: 555

Sequence: 1 atcggtggtacgaattctt.....ccagtggtgtgtaggcataag 1668

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2062474

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool p/US09712338/runat 17112003.170225 10751/app.query.fasta_1.1863
-DB=A Geneseq 19Jun03 -CFMT=fastan -SUFFIX=oli.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.gdi
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1 1 92 @runat 17112003.170225 10751 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

A Geneseq 19Jun03: *
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	554	99.8	554	19	AAW56099	A. oryzae ATCC2038
2	35	6.3	35	19	AAW57040	A. oryzae ATCC2038
3	24	4.3	551	23	ABR38864	A. niger serine ca
4	19	3.4	20	19	AAW56100	A. oryzae ATCC2038
5	16	2.9	29	19	AAW57041	A. oryzae ATCC2038
6	10	1.8	526	23	ABR38817	A. niger serine ca
7	9	1.6	144	22	AAU46145	Propionibacterium
8	9	1.6	178	22	AAU46801	Propionibacterium
9	9	1.6	454	21	AGI7034	Arabidopsis thalia
10	9	1.6	461	21	AGI7033	Arabidopsis thalia
11	9	1.6	461	21	AGI7032	Arabidopsis thalia
12	9	1.6	461	21	AGI7032	Arabidopsis thalia
13	9	1.6	473	21	AGI7032	Arabidopsis thalia
14	9	1.6	473	21	AGI7032	Arabidopsis thalia
15	9	1.6	473	21	AGI7032	Arabidopsis thalia
16	9	1.6	508	23	ABR38859	Arabidopsis thalia
17	9	1.6	554	23	ABR38859	Arabidopsis thalia
18	9	1.6	623	23	ABR38865	S. cerevisiae BAX-
19	9	1.6	1513	22	ABR38865	A. niger serine ca
20	8	1.4	69	22	AAU51334	Drosophila melanog
21	8	1.4	78	21	ABR38865	Propionibacterium
22	8	1.4	79	23	ABP06677	Zea mays protein f
23	8	1.5	81	22	AAU45980	Human ORFX protein
24	8	1.4	81	24	ABU07821	Propionibacterium
25	8	1.4	91	22	AAU34418	Corn C-4 methyl st
26	8	1.4	91	22	AAU37071	Staphylococcus aur
27	8	1.4	91	22	AAU37371	Staphylococcus aur
28	8	1.4	91	22	AAU37567	Staphylococcus aur
29	8	1.4	93	22	AAU07036	Human polypeptide
30	8	1.4	102	23	ABU60975	Human polypeptide
31	8	1.4	112	22	ABG21799	Lung specific prot
32	8	1.4	116	23	ABR03424	Novel human diagno
33	8	1.4	128	21	ABG35382	Human ORFX protein
34	8	1.4	153	22	ABG35330	Zea mays protein f
35	8	1.4	153	22	ABR38670	Human liver peptid
36	8	1.4	153	22	ABR23746	Peptide #6176 enco
37	8	1.4	153	22	AAW59299	Protein #5745 enco
38	8	1.4	153	22	AAW71846	Human brain expres
39	8	1.4	153	22	AAW19335	Human bone marrow
40	8	1.4	153	22	AAW32131	Peptide #5769 enco
41	8	1.4	153	23	ABG41659	Peptide #6168 enco
42	8	1.5	210	22	ABG10691	Human peptide enco
43	8	1.4	216	15	AAW72106	Novel human diagno
44	8	1.4	221	21	AAW33367	Bovine growth horm
45	8	1.4	240	21	AAW16761	Zea mays protein f
						Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAW56099

ID AAW56099 standard; Protein; 554 AA.

XX AAW56099;

XX AC

XX 27-AUG-1998 (first entry)

DT	DT					
XX	XX					
DE	A. oryzae ATCC20386	carboxypeptidase I protein.				
XX	Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;					
KW	food industry.					
XX	Aspergillus oryzae.					
OS	Aspergillus oryzae.					
XX	Key	Location/Qualifiers				
FT	Peptide	1..18				
FT	FT	/label= signal				

FT	Protein	19..554
XX	/label= carboxypeptidase I	
XX	WO9814599-A1.	
PX		
PD	09-APR-1998	
PF	03-OCT-1997;	97WO-US1977.
PR	27-NOV-1996;	96US-0757534.
PR	04-OCT-1996;	96US-0726880.
XX	(NOVO) NOVO NORDISK BIOTECH INC.	
PA	(NOVO) NOVO-NORDISK AS.	
XX	Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly B;	
PI	Klotz A, Mathisen TE, Rey M;	
XX	WPI; 1998-240098/21.	
DR	N-PSDB; AAV28620.	
XX	Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous	
PT	substrates, useful for improving flavour of foods	
XX	Claim 1; Fig 3; 82pp; English.	
PS	This sequence represents carboxypeptidase I from Aspergillus oryzae.	
CC	This polypeptide has an optimal activity in the range of pH 3.0-7.5 at	
CC	25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4, and a	
CC	residual activity of at least 65.5% after 30 minutes at pH 4.0 and	
CC	60 deg. C. It also has the capacity to hydrolyse X from N-CBZ-Ala-X where	
CC	N-CBZ is N-carbobenzoxy and X is any amino acid. The carboxypeptidases	
CC	can be used for obtaining hydrolysates (which can be enriched in free	
CC	glutamic acid or peptide bound glutamic acid residues) from proteinaceous	
CC	substrates. The carboxypeptidases can be used in flavour-improving	
CC	compositions in the food industry. The products can also be used for the	
CC	production of polypeptides free of carboxypeptidase activity.	
XX	SQ Sequence 554 AA;	
Alignment Scores:		
Pred. No.:	0	554
Score:	554.00	554
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	99.82%	Indels: 0
DB:	19	Gaps: 0
US-09-712-338-1 (1-1668) x AAW56099 (1-554)		
QY	1 ATGGGTGGCTACGAATTTCTCAGTGTACTCCCTTGTTGAGCGAGTGGGCCCTTCCA	60
Db	1 MetArgGlyTy rGluPheLeuSerValLeuProLeuValAlaA laSerTrpAlaLeuPro	20
QY	61 GGAAGTACACGGCGTCGTCGCTAGAGACAGCTACCACAGAACCCACCGGGSTCAAG	120
Db	21 GlySerThrProAlaSerValGlyArgGlnLeuProLysAsnProThrGlyVallys	40
QY	121 ACTCTTACAACCGCAAAACAATGTCCACATCGGTACAAGAACCCGGGGCAGAGGCCTC	180
Db	41 ThrLeuThrThrAlaAsnValThrIleArgTy rLysGluProGlyAlaGluGlyVal	60
QY	181 TCGGAGACTACCCGGGTGTCMAATCCTACTCTGATATGCGACACCTCTCCGAGTCC	240
Db	61 CysGluThrThrProGlyVallysSerTy rSerGlyTy rValAspThrSerProGluSer	80
QY	241 CATACCTCTCTCTGCTTCTCGAAGCCAGACATAACCCAGAAACTGCACCTATCACATTG	300
Db	81 HisThrPhePheTipPheGluAlaargHisAsnP roGluThrAlaProileThrLeu	100
QY	301 TGGTTGAATGGTGGCCCTGGAAAGCATTTCTTGATCGGTCTCTTCGAAAGATTGGGCCCT	360
Db	101 TrpLeuAsnGlyVGLProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyPro	120

QY 1441 AATTCTCTTCACTCCGCTATGAGCGGCGCATGAAGTCCATACCTACCGCCATC 1500
 Db |||||
 481 AsnPhSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle 500
 QY 1501 GCCTCCCTGCAATTGTTTACCGGACTATCTTCGTTGGGATATCGAGAGGCCAGAG 1560
 Db |||||
 501 AlaSerLeuGlnLeuPheAsnA-gThrIlePheGlyTyrAspIleAlaGluGlyGlnLys 520
 QY 1561 AGATCTGGCCAGCTACAGAGCAATGGAACGCTACAGCTACGCATACACAGTCTGTC 1620
 Db |||||
 521 LysIleTyrProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer 540
 QY 1621 GTGCCCTGCCTACCGCTACCGCATGTCAGTGTGGTATG 1662
 Db |||||
 541 ValProLeuProThrAlaThrSerMetSerSerValGlyMet 554

RESULT 2

AAW57040

ID AAW57040 standard; Protein; 35 AA.

XX AC AAW57040;

XX DT 27-AUG-1998 (first entry)

XX DE A. oryzae ATCC20386 carboxypeptidase I peptide fragment #3.

XX KW Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous; food industry.

XX OS Aspergillus oryzae.

XX PN WO9814599-A1.

XX PD 09-APR-1998.

XX PF 03-OCT-1997; 97WO-US17977.

XX PR 27-NOV-1996; 96US-0757534.

XX PR 04-OCT-1996; 96US-0726880.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;

XX PI Klotz A, Mathisen TE, Rey M;

XX DR WPI; 1998-240098/21.

XX PT Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous substrates, useful for improving flavour of foods

XX PS Example 2; Page 42; 82pp; English.

XX CC This sequence represents a peptide fragment from carboxypeptidase I isolated from Aspergillus oryzae and generated from cyanogen bromide cleavage. This polypeptide has an optimal activity in the range of pH 3.0-7.5 at 25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4, and a residual activity of at least 65.5% after 30 minutes at pH 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from N-CBZ-Ala-X where N-CBZ is N-carboxy and X is any amino acid. The carboxypeptidases can be used for obtaining hydrolysates (which can be enriched in free glutamic acid or peptide bound glutamic acid residues) from proteinaceous substrates. The carboxypeptidases can be used in flavour-improving compositions in the food industry. The products can also be used for the production of polypeptides free of carboxypeptidase activity.

XX SQ Sequence 35 AA;

Alignment Scores:

Pred. No.: 1.14e-25 Length: 35

Score: 35.00 Matches: 35

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.31% Indels: 0
 DB: 19 Gaps: 0

US-09-712-338-1 (1-1668) x AAW57040 (1-35)

QY 1147 GAGCTATCGGCTCAACATCACTACACCCAGTCCAATATGAGCTTACTAGCTTTC 1206
 Db 1 ASPAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyrAlaPhe 20

QY 1207 CAGCAACAGGCGACTTGTCTGGCCCACTTCATCGAAGACCTC 1251
 Db |||||

21 GlnGlnThrGlyAspPheValTyrProAsnPheIleGluAspLeu 35

RESULT 3

ABR38864

ID ABR38864 standard; Protein; 551 AA.

XX AC ABR38864;

XX DT 24-APR-2003 (first entry)

XX DE A. niger serine carboxypeptidase polypeptide #3.

XX KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW Protein solubility; viscosity; taste; texture; nutritional value;
 XX EC3.4.16.6.

XX OS Aspergillus niger.

XX PN WO200268623-A2.

XX PD 06-SEP-2002.

XX PF 22-FEB-2002; 2002WO-EP01984.

XX PR 23-FEB-2001; 2001EP-0200657.

XX PR 23-FEB-2001; 2001EP-0200658.

XX PR 23-FEB-2001; 2001EP-0200660.

XX PR 26-FEB-2001; 2001EP-0200706.

XX PR 26-FEB-2001; 2001EP-0200707.

XX PR 26-FEB-2001; 2001EP-0200708.

XX PR 26-FEB-2001; 2001EP-0200719.

XX PR 28-MAR-2001; 2001EP-0000075.

XX PR 28-MAR-2001; 2001EP-0000078.

XX PR 28-MAR-2001; 2001EP-0000080.

XX PR 28-MAR-2001; 2001EP-0000087.

XX PR 28-MAR-2001; 2001EP-0000088.

XX PR 21-MAY-2001; 2001EP-0000156.

XX PR 21-MAY-2001; 2001EP-0000159.

XX PR 21-MAY-2001; 2001EP-0000160.

XX PR 21-MAY-2001; 2001EP-0000162.

XX PR 21-MAY-2001; 2001EP-0000165.

XX PR 21-MAY-2001; 2001EP-0000166.

XX PR 21-MAY-2001; 2001EP-0000168.

XX PR 21-JUN-2001; 2001EP-0000240.

XX PR 21-JUN-2001; 2001EP-0000242.

XX PR 21-JUN-2001; 2001EP-0000244.

XX PR 12-JUL-2001; 2001EP-0000246.

XX PR 12-JUL-2001; 2001EP-0000280.

XX PR 30-JUL-2001; 2001EP-0000285.

XX PR 30-JUL-2001; 2001EP-0000323.

XX PR 30-JUL-2001; 2001EP-0000327.

XX PR 02-AUG-2001; 2001EP-0000341.

XX PR 02-AUG-2001; 2001EP-0000342.

XX PR 02-AUG-2001; 2001EP-0000343.

XX PR 02-AUG-2001; 2001EP-0000344.

XX PR 09-AUG-2001; 2001EP-0000357.

XX PR 16-AUG-2001; 2001EP-0000374.

XX PR 16-AUG-2001; 2001EP-0000377.

XX PR 20-SEP-2001; 2001EP-0000478.

XX PR 20-SEP-2001; 2001EP-0000483.

XX PR 22-OCT-2001; 2001EP-0000552.

CC manufacture of biological detergents. A polypeptide may also be useful
CC for improving protein solubility, extraction yields, viscosity or taste,
CC texture, nutritional value, minimising of antigenicity or
CC anti-nutritional factors, colour or functionality as well as processing
CC aspects like filterability of the proteinaceous raw material. The
CC sequences shown in ABR38313-ABR3869 represent the A. niger proteases of
CC the invention.

XX Sequence 526 AA;
SQ Alignment Scores:
Pred. No.: 0.808 Length: 526
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 23 Gaps: 0

US-09-712-338-1 (1-1668) x ABR38817 (1-526)

QY 679 ACGGAGCTATGGAGGCACATGCTCT 708
DB 181 ThrGluSerTyrGlyGlyHisTyrGlyPro 190

RESULT 7
AAU46145
ID AAU46145 standard; Protein; 144 AA.

XX AC AAU46145;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #7041.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59529.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX Example 1; SEQ ID No 7340; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 144 AA;
SQ Alignment Scores:

Pred. No.: 9.68 Length: 144
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 22 Gaps: 0

US-09-712-338-1 (1-1668) x AAU46145 (1-144)

QY 1224 TGCTGGCCCACTTCATCGAGACCT 1250
DB 15 CysLeuAlaGlnLeuHisArgPro 23

RESULT 8
AAU46801

ID AAU46801 standard; Protein; 178 AA.

XX AC AAU46801;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #7697.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59535.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX Example 1; SEQ ID No 7996; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 178 AA;

Alignment Scores:
 Pred. No.: 9.4 Length: 178
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 22 Gaps: 0

US-09-712-338-1 (1-1668) x AAU46801 (1-178)

QY 16 TTTCTCTCAGTCTACCTGGTGGCA 42

Db 151 PheLeuSerValLeuProLeuValAla 159

RESULT 9

AAG17034

ID AAG17034 standard; Protein; 454 AA.

AC AAG17034;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 17902.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 23-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0127462.

PR 16-APR-1999; 99US-0128234.

PR 19-APR-1999; 99US-0128714.

PR 21-APR-1999; 99US-0129845.

PR 23-APR-1999; 99US-0130077.

PR 28-APR-1999; 99US-0130449.

PR 30-APR-1999; 99US-0130510.

PR 30-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.

```
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145102.
PR 23-JUL-1999; 99US-0145118.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.26 Length: 454
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 21 Gaps: 0

US-09-712-338-1 (1-1668) x AAG17034 (1-454)
QY 295 ACATTGGTTCATGTCGGCCCTGGA 321
Db 59 ThrLeuTrpLeuAsnGlyGlyProGly 67

RESULT 10
AAG17033
ID AAG17033 standard; Protein; 461 AA.
XX AAG17033;
AC AAG17033;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 17901.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX OS
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
```

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 08-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

```
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.24
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.62%
DB: 21
Gaps: 0

US-09-712-338-1 (1-1668) x AAG17033 (1-461)
QY 295 ACATGTGGTGAATGAGTGGCCCTGGA 321
DB 66 ThrLeuTrpLeuAsnGlyGlyProGly 74

RESULT 11
AAG49759
ID AAG49759 standard; Protein; 461 AA.
XX
AC AAG49759;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62981.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
```

```
PR 23-JUL-1999; 99US-01451145.
PR 23-JUL-1999; 99US-01452218.
PR 23-JUL-1999; 99US-01452224.
PR 26-JUL-1999; 99US-01452276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.24 Length: 461
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 21 Gaps: 0

US-09-712-338-1 (1-1668) x AAG49759 (1-461)
QY 295 ACATTGCTGTTGAATGCTGCCCTGGA 321
Db 66 ThrLeuTriLeuAsnGlyGlyProGly 74

RESULT 12
AAG49762
ID AAG49762 standard; Protein; 461 AA.
XX
AC AAG49762;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62985.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
```

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139113.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 8.24 Length: 461
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 21 Gaps: 0

US-09-712-338-1 (1-1668) x AAG49762 (1-461)

QY 295 ACATTGGTGAATGGTCCCTGGA 321

Db 66 ThrLeuTrpLeuAsnGlyGlyProGly 74

RESULT 13

AAGI7032

ID AAGI7032 standard; Protein; 473 AA.

XX AC AAGI7032;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 17900.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.

```

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156558.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.21 Length: 473
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 21 Gaps: 0

US-09-712-338-1 (1-1668) x AAG17032 (1-473)
QY 295 ACATTGGTGAATGGTGGCCCTGGA 321
Db 78 ThrLeuTrpLeuAsnGlyGlyProGly 86

RESULT 14
AAG49758
ID AAG49758 standard; Protein; 473 AA.
XX AAG49758;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62980.
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2008EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
```


PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135363.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

[illegible]

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149302.
PR 23-AUG-1999; 99US-0149330.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.21      Length: 473
Score: 9.00          Matches: 9
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.62%      Indels: 0
DB: 21                Gaps: 0

US-09-712-338-1 (1-1668) x AAG49761 (1-473)
QY 295 ACATTGTGGTTGAATGGTGGCCCTGGA 321
   |||||
Db 78 ThrLeuTrpLeuAsnGlyGlyProGly 86

Search completed: November 21, 2003, 17:51:01
Job time : 59 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:49:25 ; Search time 19.5 Seconds
(without alignments)
7238.411 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 555
Sequence: 1 atgcgtggtaacgaattct.....ccagtgtgtatggcatag 1668

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Word size: 1

Total number of hits satisfying chosen parameters: 557304

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=oli.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1 1 31 @runat 17112003 170227 10817 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	99.8	554	US-08-943-714-2	Sequence 2, Appli
2	35	6.3	35	US-08-943-714-5	Sequence 5, Appli
3	19	3.4	20	US-08-943-714-3	Sequence 3, Appli
4	16	2.9	29	US-08-943-714-6	Sequence 6, Appli
5	10	1.8	423	US-08-943-714-9	Sequence 9, Appli
6	8	1.4	81	US-09-439-554-2	Sequence 2, Appli
7	8	1.4	164	US-09-252-991A-20615	Sequence 20615, A
8	8	1.4	164	US-09-252-991A-30154	Sequence 30154, A
9	8	1.4	283	US-09-107-532A-5418	Sequence 5418, Ap
10	8	1.4	291	US-09-439-554-20	Sequence 20, Appli
11	8	1.5	386	US-09-252-991A-27381	Sequence 27381, A
12	8	1.4	421	US-08-807-263-4	Sequence 4, Appli

13	1.4	8	1.4	480	2	US-08-828-488-8	Sequence 8, Appli
14	1.4	8	1.4	480	4	US-09-299-689A-8	Sequence 8, Appli
15	1.4	8	1.4	480	4	US-09-702-705-336	Sequence 336, App
16	1.4	8	1.4	480	4	US-09-736-457-336	Sequence 336, App
17	1.4	8	1.4	491	1	US-09-640-305-4	Sequence 4, Appli
18	1.4	8	1.4	491	1	US-08-360-673-4	Sequence 4, Appli
19	1.4	8	1.4	532	2	US-08-899-324-33	Sequence 33, Appli
20	1.4	8	1.4	532	3	US-08-329-892B-33	Sequence 33, Appli
21	1.4	8	1.4	557	1	US-08-309-341-2	Sequence 2, Appli
22	1.4	8	1.4	557	1	US-08-309-341-2	Sequence 2, Appli
23	1.4	8	1.4	557	1	US-08-608-267-2	Sequence 4, Appli
24	1.4	8	1.4	557	1	US-08-608-267-4	Sequence 2, Appli
25	1.4	8	1.4	557	1	US-08-608-452-2	Sequence 2, Appli
26	1.4	8	1.4	557	1	US-08-608-452-2	Sequence 2, Appli
27	1.4	8	1.4	557	1	US-08-608-224-2	Sequence 2, Appli
28	1.4	8	1.4	557	1	US-08-608-224-4	Sequence 2, Appli
29	1.4	8	1.4	557	2	US-08-967-149-2	Sequence 2, Appli
30	1.4	8	1.4	557	2	US-08-967-149-4	Sequence 4, Appli
31	1.4	8	1.4	654	4	US-09-252-991A-23035	Sequence 23035, A
32	1.4	8	1.4	679	4	US-09-252-991A-25400	Sequence 25400, A
33	1.3	9	1.3	9	3	US-08-159-339A-528	Sequence 528, App
34	1.3	10	1.3	10	3	US-08-159-339A-549	Sequence 549, App
35	1.3	19	1.3	19	3	US-08-943-714-4	Sequence 4, Appli
36	1.3	34	1.3	34	3	US-08-810-009-24	Sequence 24, Appli
37	1.3	55	1.3	55	2	US-08-637-759B-348	Sequence 348, App
38	1.3	55	1.3	55	3	US-08-871-355A-348	Sequence 348, App
39	1.3	55	1.3	55	4	US-09-201-945-348	Sequence 348, App
40	1.3	56	1.3	56	3	US-08-866-545-10	Sequence 10, Appli
41	1.3	78	1.3	78	2	US-08-831-575-5	Sequence 5, Appli
42	1.3	83	1.3	83	4	US-09-252-991A-19155	Sequence 19155, A
43	1.3	85	1.3	85	3	US-08-894-173-54	Sequence 54, Appli
44	1.3	85	1.3	85	3	US-08-894-173-55	Sequence 55, Appli
45	1.3	85	1.3	85	3	US-08-894-173-56	Sequence 56, Appli

ALIGNMENTS

RESULT 1

US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:

APPLICANT: Blinkovsky, Alexander

APPLICANT: Berka, Randy

APPLICANT: Rey, Michael

APPLICANT: Golightly, Elizabeth

APPLICANT: Klotz, Alan

APPLICANT: Mathisen, Thomas Erik

APPLICANT: Dambmann, Claus

TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids

TITLE OF INVENTION: Encoding Same

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESS: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.

STREET: 405 Lexington Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,714

FILING DATE: 03-OCT-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4990,200-US

TELECOMMUNICATION INFORMATION:

RESULT 2
US-08-943-714-5
: Sequence 5, Application US/08943714
: Patent No. 6187578
: GENERAL INFORMATION:
: APPLICANT: Blinkovsky, Alexander
: APPLICANT: Berka, Randy
: APPLICANT: Rey, Michael
: APPLICANT: Golighly, Elizabeth
: APPLICANT: Klotz, Alan
: APPLICANT: Mathisen, Thomas Erik

APPLICANT: Dambmann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 61875780 No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6187578e
US-08-943-714-5

Alignment Scores:
Pred. No.: 1.65e-26 Length: 35
Score: 35.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.31% Indels: 0
DB: 3 Gaps: 0

US-09-712-338-1 (1-1668) x US-08-943-714-5 (1-35)
QY 1147 GACGCTATCGGCTCAACATCACTACACCCAGTCCAAATAGCTCTACTAGCTTTC 1206
Db 1 AspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAspValTyrAlaPhe 20
QY 1207 CAGCAACAGGCGACTTTCTCTGGCCCACTTCATCGAGACCTC 1251
Db 21 GlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeu 35

RESULT 3
US-08-943-714-3
Sequence 3, Application US/08943714
Patent No. 6187578
GENERAL INFORMATION:
APPLICANT: Blinkovsky, Alexander
APPLICANT: Berka, Randy
APPLICANT: Rey, Michael
APPLICANT: Golightly, Elizabeth
APPLICANT: Klotz, Alan
APPLICANT: Mathisen, Thomas Erik
APPLICANT: Dambmann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 61875780 No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue

CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6187578e
US-08-943-714-3

Alignment Scores:
Pred. No.: 2.22e-10 Length: 20
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.42% Indels: 0
DB: 3 Gaps: 0

US-09-712-338-1 (1-1668) x US-08-943-714-3 (1-20)
QY 688 TATGAGGCGCACTATGCTCTGCATCTTCAATCATTTTACGACGAGATGAGAGA 744
Db 2 TyrGlyGlyHisTyrGlyProAlaPheAsnHisPheTyrGluGlnAsnGluArg 20

RESULT 4
US-08-943-714-6
Sequence 6, Application US/08943714
Patent No. 6187578
GENERAL INFORMATION:
APPLICANT: Blinkovsky, Alexander
APPLICANT: Berka, Randy
APPLICANT: Rey, Michael
APPLICANT: Golightly, Elizabeth
APPLICANT: Klotz, Alan
APPLICANT: Mathisen, Thomas Erik
APPLICANT: Dambmann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 61875780 No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714

us-09-712-338-1-oli.ra

Mon Nov 24 13:41:00 2003

FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6187578e
US-08-943-714-6
Alignment Scores:
Pred. No.: 2,27e-07 Length: 29
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.88% Indels: 0
DB: 3 Gaps: 0
US-09-712-338-1 (1-1668) x US-08-943-714-6 (1-29)
QY 1045 TAGCGCTTGTGTCGTCGTGTATGATATTCGGCATCCATATGAT 1092
Db 11 TyrAlaPheAlaGlyArgGlyValTyraSpIleArgHisProTyraSp 26
RESULT 5
US-08-943-714-9
Sequence 9, Application US/08943714
Patent No. 6187578
GENERAL INFORMATION:
APPLICANT: Blinkovsky, Alexander
APPLICANT: Berka, Randy
APPLICANT: Rey, Michael
APPLICANT: Golightly, Elizabeth
APPLICANT: Klotz, Alan
APPLICANT: Mathisen, Thomas Erik
APPLICANT: Dammann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-943-714-9
Alignment Scores:
Pred. No.: 0.21 Length: 423
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 3 Gaps: 0
US-09-712-338-1 (1-1668) x US-08-943-714-9 (1-423)
QY 679 ACGGAGCTATGAGGCGCATATGTCCT 708
Db 141 ThrGluSerTyTGlyGlyHisTyGlyPro 150
RESULT 6
US-09-439-554-2
Sequence 2, Application US/09439554
Patent No. 6479733
GENERAL INFORMATION:
APPLICANT: Rafalski, Jan Antoni
APPLICANT: Odell, Joan T.
APPLICANT: Sakai, Hajime
APPLICANT: Thorpe, Catherine J.
APPLICANT: Kinney, Anthony J.
APPLICANT: Famodu, Omolayo O.
TITLE OF INVENTION: STEROL METABOLISM ENZYMES
FILE REFERENCE: BB1114 US NA
CURRENT APPLICATION NUMBER: US/09/439,554
CURRENT FILING DATE: 1999-11-12
EARLIER APPLICATION NUMBER: 60/108,351
EARLIER FILING DATE: 1998-No. 6479733ember-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 81
TYPE: PRT
ORGANISM: Zea mays
US-09-439-554-2
Alignment Scores:
Pred. No.: 24.2 Length: 81
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 4 Gaps: 0
US-09-712-338-1 (1-1668) x US-09-439-554-2 (1-81)
QY 619 TTCCTTAGTGGACTACCTAGCTTG 642
Db 29 PheLeuSerGlyLeuProSerLeu 36
RESULT 7
US-09-252-991A-20615
Sequence 20615, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20615
 ; LENGTH: 164
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-20615

Alignment Scores:
 Pred. No.: 23.1 Length: 164
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.44%
 DB: 4
 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-252-991A-20615 (1-164)

QY 1559 AGAAGATCTGGCCAGCTACAAGA 1582
 |||||
 Db 78 ArgArgSerGlyProAlaThrArg 85

RESULT 8

US-09-252-991A-30154
 ; Sequence 30154, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30154
 ; LENGTH: 164
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-30154

Alignment Scores:
 Pred. No.: 23.1 Length: 164
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.44%
 DB: 4
 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-252-991A-30154 (1-164)

QY 1559 AGAAGATCTGGCCAGCTACAAGA 1582
 |||||
 Db 78 ArgArgSerGlyProAlaThrArg 85

RESULT 9

US-09-107-532A-5418
 ; Sequence 5418, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts

COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 5418:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 283 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...283
 SEQUENCE DESCRIPTION: SEQ ID NO: 5418:
 US-09-107-532A-5418

Alignment Scores:
 Pred. No.: 22.3 Length: 283
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.44%
 DB: 4
 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-107-532A-5418 (1-283)

QY 1135 GACTCTGTCATGGAGCTATGGC 1158
 |||||
 Db 187 AspSerValMetAspAlaIleGly 194

RESULT 10

US-09-439-554-20
 ; Sequence 20, Application US/09439554
 ; Patent No. 6479733
 ; GENERAL INFORMATION:
 ; APPLICANT: Rafalski, Jan Antoni
 ; APPLICANT: Odeil, Joan T.
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Thorpe, Catherine J.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Pamodu, Omolayo O.
 ; TITLE OF INVENTION: STEROL METABOLISM ENZYMES
 ; FILE REFERENCE: BB114 US NA
 ; CURRENT APPLICATION NUMBER: US/09/439,554
 ; CURRENT FILING DATE: 1999-11-12
 ; EARLIER APPLICATION NUMBER: 60/108,351
 ; EARLIER FILING DATE: 1998-No. 6479733ember-13
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 20
 ; LENGTH: 291
 ; TYPE: PRT

```
; ORGANISM: Zea mays
US-09-439-554-20

Alignment Scores:
Pred. No.: 22.2 Length: 291
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-439-554-20 (1-291)
QY 619 TTCCTTAGTGACTACCTAGCTTG 642
Db 61 PheLeuSerGlyLeuProSerLeu 68

RESULT 11
US-09-252-991A-27381
; Sequence 27381, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27381
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27381

Alignment Scores:
Pred. No.: 21.8 Length: 388
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-252-991A-27381 (1-388)
QY 1587 ATTCGCTTTGTAGCTGGCCAGAT 1564
Db 326 IleArgLeuValAlaGlyProasp 333

RESULT 12
US-08-807-263-4
; Sequence 4, Application US/08807263C
; Patent No. 5985627
; GENERAL INFORMATION:
; APPLICANT: Mortensen, Uffe
; APPLICANT: Olesen, Kjeld
; APPLICANT: Stennicke, Henning
; APPLICANT: Sorensen, Steen B.
; APPLICANT: Breddam, Klaus
; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
; FILE REFERENCE: 8648.71us01-no4
; CURRENT APPLICATION NUMBER: US/08/807,263C
; CURRENT FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
```

```
; ORGANISM: Zea mays
US-09-439-554-20

Alignment Scores:
Pred. No.: 21.7 Length: 421
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x US-08-807-263-4 (1-421)
QY 298 TTGTGTTGATGCTGGCCCTGGA 321
Db 48 LeuTrpLeuAsnGlyGlyProGly 55

RESULT 13
US-08-828-488-8
; Sequence 8, Application US/08828488
; Patent No. 5925521
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,488
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 190283
; US-08-828-488-8

Alignment Scores:
Pred. No.: 21.5 Length: 480
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0
```

US-09-712-338-1 (1-1668) x US-09-828-488-8 (1-480)

Qy 298 TTGTGTTGAATGGTGGCCCTGGA 321
Db 80 LeuTrpLeuAsnGlyGlyProgly 87

RESULT 14

US-09-299-689A-8

; Sequence 8, Application US/09299689A
; Patent No. 6379913

GENERAL INFORMATION:

; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,689A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,488
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

SEQUENCE CHARACTERISTICS:

; LENGTH: 480 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 190283

US-09-299-689A-8

Alignment Scores:

Pred. No.: 21.5

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.44%

DB: 4

Length: 480

Matches: 8

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-712-338-1 (1-1668) x US-09-299-689A-8 (1-480)

Qy 298 TTGTGTTGAATGGTGGCCCTGGA 321

Db 80 LeuTrpLeuAsnGlyGlyProgly 87

RESULT 15

US-09-702-705-336

; Sequence 336, Application US/09702705

; Patent No. 6504010

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.478C14

; CURRENT APPLICATION NUMBER: US/09/702,705

; NUMBER OF SEQ ID NOS: 1833

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 336

; LENGTH: 480

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-702-705-336

Alignment Scores:

Pred. No.: 21.5

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.44%

DB: 4

Length: 480

Matches: 8

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-712-338-1 (1-1668) x US-09-702-705-336 (1-480)

Qy 298 TTGTGTTGAATGGTGGCCCTGGA 321

Db 80 LeuTrpLeuAsnGlyGlyProgly 87

Search completed: November 21, 2003, 17:56:40

Job time : 25.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:40:04 ; Search time 63.5 Seconds
(without alignments)
9590.842 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 3027
Sequence: 1 atcggtggctacgaattct.....ccagtggtgtgtagcatag 1668

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Zgapop 6.0 , Zgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 1332376

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODELS=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US09712338/runat_17112003_170150_10599/app_query.fasta_1.1863
-DB=Published Applications AA -OPMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USAP=US09712338 @CGN 1 1 15 @runat_17112003_170150_10599
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result	Score	Match	Length	DB	ID	Description
No.						

SUMMARIES

1	517.5	17.1	421	9	US-09-420-785A-4	Sequence 4, Appli
2	517.5	17.1	421	9	US-09-901-252-15	Sequence 15, Appli
3	361.5	11.9	476	14	US-10-084-018-3	Sequence 3, Appli
4	361.5	11.9	476	10	US-09-909-320-164	Sequence 164, App
5	360.5	11.9	476	10	US-09-909-088B-164	Sequence 164, App
6	360.5	11.9	476	10	US-09-905-291A-164	Sequence 164, App
7	360.5	11.9	476	10	US-09-902-853-164	Sequence 164, App
8	360.5	11.9	476	10	US-09-907-824-164	Sequence 164, App
9	360.5	11.9	476	10	US-09-907-841-164	Sequence 164, App
10	360.5	11.9	476	11	US-09-904-011-164	Sequence 164, App
11	360.5	11.9	476	11	US-09-906-742-164	Sequence 164, App
12	360.5	11.9	476	11	US-09-906-838-164	Sequence 164, App
13	360.5	11.9	476	11	US-09-907-613-164	Sequence 164, App
14	360.5	11.9	476	11	US-09-907-942-164	Sequence 164, App
15	360.5	11.9	476	11	US-09-796-753-40	Sequence 40, Appli
16	360.5	11.9	476	11	US-09-904-859-164	Sequence 164, App
17	360.5	11.9	476	11	US-09-909-204-164	Sequence 164, App
18	360.5	11.9	476	11	US-09-904-820-164	Sequence 164, App
19	360.5	11.9	476	11	US-09-904-786-164	Sequence 164, App
20	360.5	11.9	476	11	US-09-906-646-164	Sequence 164, App
21	360.5	11.9	476	11	US-09-906-700-164	Sequence 164, App
22	360.5	11.9	476	11	US-09-903-786-164	Sequence 164, App
23	360.5	11.9	476	11	US-09-902-903-164	Sequence 164, App
24	360.5	11.9	476	11	US-09-903-749A-164	Sequence 164, App
25	360.5	11.9	476	11	US-09-904-119-164	Sequence 164, App
26	360.5	11.9	476	11	US-09-904-956-164	Sequence 164, App
27	360.5	11.9	476	11	US-09-902-736-164	Sequence 164, App
28	360.5	11.9	476	11	US-09-907-794-164	Sequence 164, App
29	360.5	11.9	476	11	US-09-903-943-164	Sequence 164, App
30	360.5	11.9	476	11	US-09-904-462-164	Sequence 164, App
31	360.5	11.9	476	11	US-09-907-925-164	Sequence 164, App
32	360.5	11.9	476	11	US-09-902-692-164	Sequence 164, App
33	360.5	11.9	476	11	US-09-903-520-164	Sequence 164, App
34	360.5	11.9	476	11	US-09-905-056-164	Sequence 164, App
35	360.5	11.9	476	11	US-09-909-064-164	Sequence 164, App
36	360.5	11.9	476	11	US-09-904-553-164	Sequence 164, App
37	360.5	11.9	476	11	US-09-905-381-164	Sequence 164, App
38	360.5	11.9	476	11	US-09-905-088-164	Sequence 164, App
39	360.5	11.9	476	11	US-09-907-575-164	Sequence 164, App
40	360.5	11.9	476	11	US-09-905-075-164	Sequence 164, App
41	360.5	11.9	476	11	US-09-902-759-164	Sequence 164, App
42	360.5	11.9	476	11	US-09-902-634-164	Sequence 164, App
43	360.5	11.9	476	11	US-09-902-713-164	Sequence 164, App
44	360.5	11.9	476	11	US-09-907-979-164	Sequence 164, App
45	360.5	11.9	476	11	US-09-902-615-164	Sequence 164, App

ALIGNMENTS

RESULT 1

US-09-420-785A-4

; Sequence 4, Application US/09420785A

; Patent No. US20010010923A1

; GENERAL INFORMATION:

; APPLICANT: MORTENSEN, UFFE

; APPLICANT: OLESEN, KJELD

; APPLICANT: STERNICK, HENNING

; APPLICANT: SORESENSEN, STEEN B.

; APPLICANT: BREDDAM, KLAUS

; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE

; FILE REFERENCE: 089187/0109

; CURRENT APPLICATION NUMBER: US/09/420,785A

; CURRENT FILING DATE: 1999-10-19

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 421

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-420-785A-4

Alignment Scores: 2.22e-34 Length: 421

Pred. No.: 2.22e-34

Score: 517.50 Matches: 142
Percent Similarity: 43.98% Conservative: 70
Best Local Similarity: 29.46% Mismatches: 185
Query Match: 17.10% Indels: 85
DB: Gaps: 17

US-09-712-338-1 (1-1668) x US-09-420-785A-4 (1-421)

QY	151	CGGTACAAAGAACCCGGGCGAGGCGCTTCGCAGACTACC	CGCGGTGTCAAATCCTTAC	210
Db	1	LysIleLysAspProLysIleLeuGlyIle-----	AspProAsnValThrGlnTyr	17
QY	211	TCCGGATATGTGCACACTCTCCCGAG----	TCCCATACCTTTCTTGTTCTTTCGAAGCC	267
Db	18	ThrGlyTyrLeuAspValGluAaspGluAaspLysHisPhePheTrpThrPhelGluSer	37	
QY	268	AGACATAACCCAGAACAATGCACATTCACATTGTGTTCAATGTGGCCCTCGAAGCGAT	327	
Db	38	ArgAsnAspProAlaLysAspProValIleLeuTrpLeuAsnGlyGlyProGlyCysSer	57	
QY	328	TCTTTGATCGGTCTCTTCGAAAGAGTTGGGCCCTTGCCATGTC	CAATTTCGACTTTTTCATGAC	387
Db	58	SerLeuThrGlyLeuPhePheGluLeuGlyProSerSerIleGlyProAspLeuLysPro	77	
QY	388	TACATCAACCTCTACTCGTGAACGAGGTCTCCAATTTACTANTCTCTGCCAGCANTTG	447	
Db	78	IleGlyAsnProTyrSerTrpAsnSerAsnAlaThrValIlePheLeuAspGlnProVal	97	
QY	448	GGAGTCGGCTTTTCATATAGTAGATACGGTTGATGGGTGCCATTAAACCTTAACCTGGGTC	507	
Db	98	AsnValGlyPheSerTyrSer-----	104	
QY	508	GTCGAAAAATTCGAGCTTTGCGAGGAGTTTCAGGCGCGGTACCCAAACCATTGATGCCACTGTG	567	
Db	105	-----GlySerSergly-----	108	
QY	568	ATCGATACACTCAATCTTTCGCCGAGAGCGCGCTTGGAGATCTCGAAGANTTCCTTAGT	627	
Db	109	ValSerAsnThrValAlaAlaGlyLysAspValTyrAsnPheLeuGluLeuPhePheAsp	128	
QY	628	GGACTACCTAGTTGGACTCTAGGGTCAGCTCTAAGGACTTCAGTCTATGGACGGAGAGC	687	
Db	129	GlnPheProGluTyrValAsnLysGlyGln-----	AspPheHisIleAlaGlyGluSer	146
QY	688	TATGGAGGCACTATGGTCCTGCATTTCTTCAATCATTTTACGACGAGATGAGAAATT	747	
Db	147	TyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeuSerHisLysAspArg---	165	
QY	748	GCCAAACGGFAGTGTAAATGGTGTTCAGCTTAATTTCAACTCTCTGGGAATTAIT	801	
Db	166	-----AsnPheAsnLeuThrSerValLeuIleGly	175	
QY	802	AACGGCATCATCGACGAGCGGNATCCAGGCCCTTACTACCTGAATTCGCT	852	
Db	176	AsnGlyLeuThrAspProLeuThrGlnTyrAsnTyrTyrGluProMetAlaCysGlyGlu	195	
QY	853	-----GTGAACAATACTCACGGCTGTCAACGACGAGACCGTCTAC	897	
Db	196	GlyGlyGluProSerValLeuProSerGluGluCysSerAlaMetGluAspSerLeu---	214	
QY	898	AACATCATGAAGTTTGCCAAACCAATGCCAATGGTTGCCAGGATTTGATTTCCACCTGC	957	
Db	215	-----GluArgCysLeuGlyLeuIleGluSerCys	224	
QY	958	AAACACACAAACCGCACCGCATTAGTGACTACGCCCTCTCGCGGAGCCACCAACATG	1017	
Db	225	TyrAspSerGln-----SerValTrpSerCysValProAlaThrIleTyr	239	
QY	1018	TGCAGGCAACAATGTTGAGGGGCGCACTACGCGCTTTGCTGCTGGTGTGTATGATATT	1077	
Db	240	CysAsnAsnAlaGlnLeuAlaProTyrGlnArg---ThrGlyArgAsnValTyrAspile	258	
QY	1078	CGCATCTCATATGATGAC-----	CGGACTCGGCCAAGTTATTACACAAA	1122

QY 268 AGACATAACCCAGAACTGCATATACATTTGTTGAATGGTGGCCCTGGAAGCAT 327
 Db |||||:|||||
 QY 38 ArgAsnAspProAlaAlaAspProValIleLeuTrpLeuAsnGlyGlyProGlyCysSer 57
 Db |||||:|||||
 QY 328 TCTTTGATCGGTCTCTTCAAGAGTTGGCCCTGCCATGTCAATTCGACTTTTGATGAC 387
 Db |||||:|||||
 QY 58 SerLeuThrGlyLeuPheGluLeuGlyProSerSerIleGlyProAspLeuLysPro 77
 Db |||||:|||||
 QY 388 TACATCAACCTCACTCGTGGAAAGAGGTCTCCAAATTTACTATTCTCTCCAGCATTTG 447
 Db |||||:|||||
 QY 78 IleGlyAsnProTyrSerTrpAsnSerAsnAlaThrValIlePheLeuAspGlnProVal 97
 Db |||||:|||||
 QY 448 GGAGTCGGCTTTTCATATAGTATGATACAGGTGTGATGGGTCCATTAAACCTGTAACTGGGGTC 507
 Db |||||:|||||
 QY 98 AsnValGlyPheSerTyrSer----- 104
 QY 508 GTCGAAATTCGAGCTTTGAGGAGTTGAGGGCCGTACCCAACATTGATGCCACTGTG 567
 Db |||||:|||||
 QY 105 -----GlySerSerGly----- 108
 QY 568 ATCGATACCTACCAATTTGCCGACAGAGCCCTTTGGGAGATCTCTGCAAGGATTCCTTAGT 627
 Db |||||:|||||
 QY 109 ValSerAsnThrValAlaAlaGlyAspValTyrAsnPheLeuGluLeuPhePheAsp 128
 Db |||||:|||||
 QY 628 GGACTACTAGCTGGACTTAGGTCAGTCTAAGGACTTCACTCTATGTCAGCGAGAGC 687
 Db |||||:|||||
 QY 129 GlnPheProGluTyrValAsnLysGlyGln-----AspPheHisIleAlaGlyGluSer 146
 Db |||||:|||||
 QY 688 TATGGAGGCACTATGCTCGTCACTTCTCAATCATTTTACGAGCAGAAATGAGAGAATT 747
 Db |||||:|||||
 QY 147 TyrAlaGlyHisTyrIleProValPheAlaSerGluLeuSerHisLysAspArg--- 165
 QY 748 GCCAACGGTAGTTAATGGTGTTCAGCTTAATTTCAACTCTCTGGGAATTAT--- 801
 Db |||||:|||||
 QY 166 -----AspPheAsnLeuThrSerValLeuIleGly 175
 QY 802 AACGGCATATCGACGAGGCGATCCAGGCCCTTACTACCTCGAATTCGCT----- 852
 Db |||||:|||||
 QY 176 AsnGlyLeuThrAspProLeuThrGlnTyrAsnTyrTyrGluProMetAlaCysGlyGlu 195
 Db |||||:|||||
 QY 853 -----GTGACAAATACCTAGCTATCAGCTGTCAACGAGACCGTCTAC 897
 Db |||||:|||||
 QY 196 GlyGlyGluProSerValLeuProSerGluGluCysSerAlaMetGluAspSerLeu--- 214
 Db |||||:|||||
 QY 898 AACTACATGAAGTTTGCCACCAATGCCAAATGGTGGCAGGATTTGATTTCCACTGC 957
 Db |||||:|||||
 QY 215 -----GluArgCysLeuGlyLeuIleGluSerCys 224
 QY 958 AAACAGACAAACCGCAGCATTTAGCTGACTAGCCCTCTGCGCCGGAAGCCACCAACATG 1017
 Db |||||:|||||
 QY 225 TyrAspSerGln-----SerValTrpSerCysValProAlaThrIleTyr 239
 QY 1018 TGCAAGGACAAATGTTGAGGGCCATACCTAGCCCTTGTGGTGGTGGTGTGTATGATATT 1077
 Db |||||:|||||
 QY 240 CysAsnAsnAlaGlnLeuAlaProTyrGlnArg---ThrGlyArgAsnValTyrAspIle 258
 Db |||||:|||||
 QY 1078 CGCACTCCATATGATGAC-----CCGACTCCGCAAGTTATTACACAAA 1122
 Db |||||:|||||
 QY 259 ArgLysAspCysGluGlyGlyAsnLeuCysTyrProThrLeuGlnAsp---IleAspAsp 277
 Db |||||:|||||
 QY 1123 TTTCTGGCAAGAGCTCTCTATGGACGTATGGCGTCAACATCACTACACCCAGTCC 1182
 Db |||||:|||||
 QY 278 TyrLeuAsnGlnAspTyrValLysGluAlaValGlyAlaGluValAspHisTyrGluSer 297
 Db |||||:|||||
 QY 1183 ---AATAATGACGTCTACTACCTTCCAGCAAAACAGGCGACTTGTCTGGCCCACTTC 1239
 Db |||||:|||||
 QY 298 CysAsnPheAspIleAsnArgAsnPheLeuPheAlaGlyAspTrpMetLysPro---Tyr 316
 Db |||||:|||||
 QY 1240 ATCGAAGACCTCGAGGAGATCTCTCTCCCGGTGGTGTCTCTCATCTATCGCGAC 1299
 Db |||||:|||||
 QY 317 HisThrAlaValThrAspLeuLeuAsnGlnAspLeuProIleLeuValTyrAlaGlyAsp 336

QY 1300 GCCGATTACATCTGCACTGTTGCGCGGTGAGCGGTTTCCCTGCTGCGAAGTACTCC 1359
 Db |||||:|||||
 QY 337 LysAspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrAspValLeuProTrpLys 356
 Db |||||:|||||
 QY 1360 CAAGCGGCCAGTTCCGAAAGC-----GCAGGGTACAGCCCTGAAAGTCAACGCGC 1410
 Db |||||:|||||
 QY 357 TyrAspGluGluPheAlaSerGlnLysValArgAsnTrpThrAlaSerIleThrAspGlu 376
 Db |||||:|||||
 QY 1411 GTCGAGTATGGGAAACTCGCGAGTATGTAATTTCTCTTCACTGCGGTCTATGAGCA 1470
 Db |||||:|||||
 QY 377 ValAla---GlyGluValLysSerTyrLysHisPheThrTyrLeuArgValPheAsnGly 395
 Db |||||:|||||
 QY 1471 GGCCATGAGTCCCATCTACTACCGCCATCCCTCCCTGCAATTTTAAACCGGACTATC 1530
 Db |||||:|||||
 QY 396 GlyHisMetValProPheAspValProGluAsnAlaLeuSerMetValAsnGluTrpIle 415
 Db |||||:|||||
 QY 1531 TTCGGT 1536
 Db |||||:|||||
 QY 416 HisGly 417
 Db |||||:|||||
 RESULT 3
 US-10-084-018-3
 ; Sequence 3, Application US/10084018
 ; Publication No. US20020160499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hawkins, Phillip R.
 ; Hillman, Jennifer L.
 ; Lal, Preeti
 ; Goli, Survi K.
 ; TITLE OF INVENTION: NOVEL HUMAN SERINE
 ; CARBOXYPEPTIDASE
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/084,018
 ; FILING DATE: 25-Feb-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/299,689A
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: 08/828,488
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0241 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 476 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: MMLR3DT01
 ; CLONE: 566993
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-10-084-018-3

Alignment Scores:

Pred. No.:	2,41e-21	Length:	476
Score:	361.50	Matches:	141
Percent Similarity:	39.03%	Conservative:	69
Best Local Similarity:	26.21%	Mismatches:	201
Query Match:	11.94%	Indels:	128
DB:	14	Gaps:	23

US-09-712-338-1 (1-1668) x US-10-084-018-3 (1-476)			
QY	50	GGGCGCTCCAGGAAGTACACGGCGTGGTGGTAGAGACAGCTACCCAGAACCCCA	109
DB	22	GlyLeuPheHisSerLeuTyrArgSerValSerMet	Pro 34
QY	110	CGGGGTCAAGACTCTTACAAACGCAAAACAATGTACCA	160
DB	35	ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle	54
QY	161	AACCCGGGCGACAGGGCGTCTCGAGACTACCCCG	210
DB	55	GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr	74
QY	211	TCGGATAT	264
DB	75	AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPhePheThrPhePro 94	
QY	265	GCAGACATAACCCAGAACTGCACCTATCACATTTGTGTTGAATGTGGCCCTCGAAGC	324
DB	95	AlaGlnIleGlnProGluAspAlaProValValLeuTriPleuGlnGlyGlyProGlyGly	114
QY	325	GATCTTTTCGCTCTTCGAGAGTGGGGCTTGGCATGTCCATTCGACTTTTGAT	384
DB	115	SerSerMet**GlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr	134
QY	385	GACTACATCAACCTCACTCGGACGAGGTCTCCAAATTTACTATTCTGTCTCCAGCA	444
DB	135	LeuArgAspAspPheProThrThrThr***SerMetLeuTyrIleAspAsnPro	154
QY	445	TTGGGAGTCGGCTTTTCATATAGTGATACGTTGATGGFPCCATTAAACCTGTAACTGGG	504
DB	155	ValGlyThrGlyPheSerPheThrAspAspThrHisGly	167
QY	505	GTCTCGAAATTCGAGCTTTGCGAGGTTTCAGGGCTCAGGGCGGTACCAACCATGTGCACT	564
DB	168	Tyr	168
QY	565	CTGATCGATACCAATCTTCCGCGAGAGCCGCTTGGGAGATCTCGCAAGGATTCCTT	624
DB	169	AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe	187
QY	625	AGTGGACTACCTAGCTTGGACTCTAGGGTCAGTCTAAGCACTTCACTCTATGGAACGAG	684
DB	188	GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu	203
QY	685	AGTATGGAGGCACTATGCTCTGCAATCTTCAATCATTTTACGAGCAGATGAGAGA	744
DB	204	SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuLeuHisSerLeuAsn	221
QY	745	ATTGCCAACGCTAGTGTAAATGTTGTTGAGCTTAAATTTCACTCTCTGGGAATTTATAAC	804
DB	222	ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp	237
QY	805	GGCATATCGACGAGGCGATCCAGGCCCTTACTACCTGAAATTCGCTGTGAAACAATACC	864
DB	238	GlyTyrSerAspProGluSerIleIleGlyTyrAlaGluPhe	253
QY	865	TACGGTATCAGGCTGTCAACGACCGCTCTACAACTACATGAGATTTCGCCAACCAATG	924
DB	254	TyrGlnIleGlyLeuLeuAspGluLysGlnLysLysTyr	270
QY	925	CCAAATGTTCCAGGATTTGATTTCCACCTTCGCAACACAGACAAAAC	975
DB	271	CysHisGluCysIleGluHisIleArgLysGlnAsnTyrPheGluAlaPhe	287

QY	976	GCATTAGCTAGTACGCTCTCTCGCGGAGCAACCCACAC	1029
DB	288	GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn	307
QY	1030	GTGAGGGG	1083
DB	308	ValThrGlyCysSerAsnTyrTyrAsnPheLeu	318
QY	1084	CCATATGATGACCGGACTCCCGCAAGTTATTACAAATTTCTGCGCAAGGACTCTGTC	1143
DB	319	ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal	338
QY	1144	ATGACGCTATCGGCTCAACATCACTACACCCAGTCCCAATATACGCTCTAC	1197
DB	339	ArgGlnAlaIleHisValGly	357
QY	1198	TACGCTTTCCAGCAACACAGCGGACTTGTCTGGCCCAACTTCATCGAAGACTCGAGGAG	1257
DB	358	TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn	377
QY	1258	ATCCTTGCTCTCCCGTGGTGTCTCCCTCATCTATGGCGACCGGATTCATC	1311
DB	378	TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla	392
QY	1312	TGCAACTGTTCTCGGCGGTTCAGGCGGTTTCC	1341
DB	393	AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTyrLysGlySerGln	409
QY	1342	CTCGCTGCGAACTACTCCCAAGCCGCCAC	1380
DB	410	GluTyrLysLysAlaGluLysValTyrLysIlePheLysSerAspSer	426
QY	1381	GCAGGTACAGCCCTGAAAGTCAACGGGCTCGATATGGGAACTCCGAG	1434
DB	427	GluValAlaGlyTyr	434
QY	1435	TATGTTAATTTCTCTTCACTCGCGTCTATAGGAGGCGCATGAAGTCCCATACTACAG	1494
DB	435	ValGlyAspPheHisGlnValIleIleArgGlyGlyHisIleLeuProTyrAspGln	454
QY	1495	CCATCGCTCTCCCTGCAATTTGTTAAACCGGACTATCTTC	1542
DB	455	ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp	472

RESULT 4
 US-09-309-320-164
 ; Sequence 164, Application US/09909320
 ; Patent No. US20020132240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kijavlin, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/309,320
 ; CURRENT FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 60/143,048
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: US 60/145,698
 ; PRIOR FILING DATE: 1999-07-26
 ; PRIOR APPLICATION NUMBER: US 60/146,222
 ; PRIOR FILING DATE: 1999-07-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219
 ; PRIOR FILING DATE: 2000-01-05
 ; NUMBER OF SEQ ID NOS: 423
 ; SEQ ID NO 164
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-909-320-164

Alignment Scores:
 Pred. No.: 2,92e-21 Length: 476
 Score: 360.50 Matches: 141
 Percent Similarity: 39.22% Conservative: 70
 Best Local Similarity: 26.21% Mismatches: 200
 Query Match: 11.91% Indels: 128
 DB: 10 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-909-320-164 (1-476)

QY	50	GGGCGCTCCAGGAAGTACACGGCGTCCGTCGGTAGAAGACAGCTACCCAGAACCCCA	109
Db	22	GlyLeuPheArgSerLeuTyrArgSerValSerMet	34
QY	110	CGGGGCTCAAGACTTTACACCGCAACAAATGTACCA	160
Db	35	ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle	54
QY	161	AACCCGGCGCAGGGCGTCTCGGAGCTACCCCG	210
Db	55	GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr	74
QY	211	TCGTGATAT	264
Db	75	AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPhePheTrpPhePro	94

QY	265	GCCAGACATACCCAGAACTGCACCTATCATCTGTGGTGAATGTGGCCCTGGAAAGC	324
Db	95	AlaGlnIleGlnProGluAspAlaProValValLeuTrpLeuGlnGlyGlyProGlyGly	114
QY	325	GATTCTTTGATCGGTCTCTTCGAAGAGTTGGCCCTTGCACATGTCATTCACATTTTCAT	384
Db	115	SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr	134
QY	385	GACTACATCAACCTCACTCGTGAACGAGTCTCCCAATTTACTATTCTCTGCCAGCA	444
Db	135	LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro	154
QY	445	TTGGAGTCGGCTTTTCATATAGTGTACGGTTCATGGTCCATTAACCTGTAACTGGG	504
Db	155	ValGlyThrGlyPheSerPheThrAspAspThrHisGly	167
QY	505	GTCTGCAAAAATTCGAGCTTTCGAGGAGTTTCAGGGCGGTACCCCAACCATGTATGCCACT	564
Db	168		Tyr
QY	565	CTGATCATACTACCAATCTTCGCGAGAGCGCTTGGGAGATCTCTGCAAGGATTCTCT	624
Db	169	AlaValAsnGluAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe	187
QY	625	AGTGGACTACTAGCTTGGACTCTAGGTGCTAGTCTAAGGACTTCAGTCTATGACCGAG	684
Db	188		GlnIlePheProGluTyrLysAsnAspPheTyrValThrGlyGlu
QY	685	AGCTATGGAGGCACTATGCTCTGCAATCTTCAATCATTTTACGAGCAGAGATGACAGA	744
Db	204	SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn	221
QY	745	ATTGCCAAGGTAGTCTTAATGTTGTTCAGCTTAATTTCACTCTCTGGAATATTAAAC	804
Db	222		ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp
QY	805	GGCATCATCGACGAGCGATCCAGGCCCTTACTACCTGAATTCGCTGTGAACAATAC	864
Db	238	GlyTyrSerAspProGluSerIleIleGlyGlyTyrAlaGluPhe	253
QY	865	TACGGTATCAAGGCTGTCAACGAGACCGCTACACATACATGAGTTGCCACCAATATG	924
Db	254	TyrGlnIleGlyLeuLeuAspGluLysGlnLysLysTyr	270
QY	925	CCAAATGGTTCGAGGATTTGATTTCCACCTGCAAAACAGACAAAC	975
Db	271		CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe
QY	976	GCATTAGCTGACTACGCCCTCTGCGCGAAGCCCAAC	1029
Db	288	GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn	307
QY	1030	GTTGAGGGG	1083
Db	308	ValThrGlyCysSerAsnTyrTyrAsnLeu	318
QY	1084	CCATATGATGACCCGACTCCGCCAAGTATTACAAATAATTTCTGCAAGGACTCTGTC	1143
Db	319	ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal	338
QY	1144	ATGGAGCTATCGGCTCAACATCACTACCCAGTCCAAATGAGCTCTAC	1197
Db	339	ArgGlnAlaIleHisValGly	357
QY	1198	TACGGCTTTCCAGCAACAGGCGACTTTCTGTCGGCCCACTTCATCGAGACCTCGAGGAG	1257
Db	358	TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn	377
QY	1258	ATCCTTGCTCTCCCGTCTGCTGCTCCCTCATCTATGCGAGCGCGGATTACATC	1311
Db	378		TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla
QY	1312		TGCAACTGGTTCCGGCGGTACGGCGCTTCC

```

; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-164

Alignment Scores:
Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservative: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 10 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-909-088B-164 (1-476)
QY 50 GGGCCCTCCAGAGTAGTACACCGCGCTCGTGGTAGAGACAGTACCCAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
QY 110 CCGGGGTCAAGACTCTTACACCGCAACAACTGTACCA-----TCCGGTACAAG 160
Db 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle 54
QY 161 AACCCGGGCGAGAGGCGCTCTCGAGACTACCCCG-----GGTGTCAATCTCTAC 210
Db 55 GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
QY 211 TCTGGATAT-----GTGACACCTCTCCGAGTCCCATACCTCTCTGGTTCTTCGAA 264
Db 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePro 94
QY 265 GCCAGACATACCCAGAACTGCACCTATCACATTTGTGGTTGAATGGTGGCCCTGGAAGC 324
Db 95 AlaGlnIleGlnProGluAspAlaProValValLeuTyrLeuGlnGlyProGlyGly 114
QY 325 GATTCTTTGATCGTCTCTTCGAAGAGTTGGGCGCTTGGCCATGTCAATTCGACTTTTGTAT 384
Db 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr 134
QY 385 GACTATCATCAACCTCCTCGTGAACGAGTCTCCAAATTACTATTCTCTGTCCAGCCA 444
Db 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
QY 445 TTGGGAGTCCGCTTTTCATATAGTATACGTTGTGGTTCATTAACCCCTGTAACTGGG 504
Db 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly----- 167
QY 505 GTCGTGAAAATTCGAGCTTTGCAGAGTTTCAGGGCGGTACCCACCATGTGTCCT 564
Db 168 -----Tyr 168
QY 565 CTGATCGATACCTATCTTCGCGAGAGCGCGCTTGGGAGATCTTCGAGGATTCCTT 624
Db 169 AlaValAsnGluAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe--- 187
QY 625 AGTGGACTACCTAGCTTGGACTCTTAGGGTGCAGTCTAAGGACTTCAGTCTATGACGGAG 684

```

```

393 AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTrpLysGlySerGln----- 409
QY 1342 CTCGCTGGAAGTACTCCAGAGCGCCAG-----TTCGGAAGC----- 1380
Db 410 -----GluTyrLysLysAlaGluLysValTrpLysIlePheLysSerAspSer 426
QY 1381 -----GCAGGTCACACGCCCTGAAAGCTCAACGGCTCGAGTATGGGAAACTCGCGAG 1434
Db 427 GluValAlaGlyTyr-----IleArgGln 434
QY 1435 TATGTAATTTCTCTCTCAGCTCGCTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1494
Db 435 AlaGlyAspPheHisGlnValIleIleArgGlyGlyGlyHisIleLeuProTyrAspGln 454
QY 1495 CCCATCGCTCCCTGCAATGTTTAAACCGGACTATCTTC-----GGTTGGGAT 1542
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTrpAsp 472

RESULT 5
; Sequence 164, Application US/09909088B
; Patent No. US20020146702A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214

```

```

Db 188 -----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu 203
QY 685 AGCTATGAGGCACTATGGCTCTGCATCTTCATCATTTTACGAGCAAGATGAGA 744
Db 204 SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn----- 221
QY 745 ATTGCGCAACGGTAGTGTAAATGGTGTTCAGCTTAATTTCAACTCTCTGGGAATATTAAAC 804
Db 222 -----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp 237
QY 805 GGCATATCGACGAGGCGATCCAGCCCTTACTACCTCGATTCGCTGTGTAACAATACC 864
Db 238 GlyTyrSerAspProGluSerIleIleGlyGlyTyrAlaGluPhe-----Leu 253
QY 865 TACGGTATCAAGGCTGTCAACAGCACCGCTCACAACCTACAGTTCGCAACCAATG 924
Db 254 TyrGlnIleGlyLeuLeuAspGluLysGlnIlystYr-----PheGlnLysGln--- 270
QY 925 CCAATGGTTCAGGATTTGATTTCCACCTGCAACAGACAAC-----CGCAC 975
Db 271 -----CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe 287
QY 976 GCATTAGCTACTACGCTCTGCGCCGAGCAACCAAC-----ATGTGAGGGAACAAT 1029
Db 288 GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn 307
QY 1030 GTTAGGG-----CCATACTACGCTCTGCTGCTGTGTGTATGATATTCGGCAT 1083
Db 308 ValThrGlyCysSerAsnTyrTyrAsnPheLeu----- 318
QY 1084 CCATATGATACCGGACTCGCCCAAGTTATTACAAATTTCTGCAAAAGGACTCTGTC 1143
Db 319 ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal 338
QY 1144 ATGACGCTATCGCGTCACATCACTACACCGACTCCCAATATGACGCTAC----- 1197
Db 339 ArgGlnAlaIleHisValGly-----AsnGlnThrPheAsnAspGlyThrIleValGluLys 357
QY 1198 TAGCTTTTCCAGCAACAGCGCACTTGTCTGCGCCCACTTCATCGAAGACTCGAGGAG 1257
Db 358 TyrLeuArgGluAspThrValGlnSerValLysProTrpLeuThrGluIleMetAsnAsn 377
QY 1258 ATCCTTGCTCTCCCGTGTCTCCCTCATCTATGGCAGCGCGATTACATC----- 1311
Db 378 -----TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla 392
QY 1312 -----TGCAACTGTTCCGGCGGTGAGGCCGTGAGTATGGGAAACTCCCGAG 409
Db 393 AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTrpLysGlySerGln----- 410
QY 1342 CTGCTGTGCAACTACTCCCAAGCGCCCGAG-----TTCCGAAGC----- 1380
Db 410 -----GluTyrLysLysAlaGluLysValTrpLysIlePheLysSerAspSer 426
QY 1381 -----GAGGGTACACGCCCTCTGAAGTCAACGGCGTCAAGTATGGGAAACTCCCGAG 1434
Db 427 GluValAlaGlyTyr-----IleArgGln 434
QY 1435 TATGTTAAATTTCTCCTTCACTCGCTCTATGAGGAGGCGCATGAGTCCATACCTACACG 1494
Db 435 AlaGlyAspPheHisGlnValIleIleArgGlyGlyHisIleLeuProTyrAspGln 454
QY 1495 CCCATCGCTCCCTGCAATTTGTTTAAACCGGACTATCTTC-----GGTTGGGAT 1542
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTrpAsp 472

```

RESULT 6

US-09-905-291A-164
 ; Sequence 164, Application US/09905291A
 ; Patent No. US20020160374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, A.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth, J.
 APPLICANT: Kijavir, Ivar J.
 APPLICANT: Mather, Jennie P.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William, I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: 10466-14
 CURRENT APPLICATION NUMBER: US/09/905,291A
 CURRENT FILING DATE: 2001-07-12
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: US 60/143,048
 PRIOR FILING DATE: 1999-07-07
 PRIOR APPLICATION NUMBER: US 60/145,698
 PRIOR FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: US 60/146,222
 PRIOR FILING DATE: 1999-07-28
 PRIOR APPLICATION NUMBER: PCT/US99/21090
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/21547
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/23089
 PRIOR FILING DATE: 1999-10-05
 PRIOR APPLICATION NUMBER: PCT/US99/28214
 PRIOR FILING DATE: 1999-11-29
 PRIOR APPLICATION NUMBER: PCT/US99/28313
 PRIOR FILING DATE: 1999-11-30
 PRIOR APPLICATION NUMBER: PCT/US99/28564
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/28565
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/30095
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: PCT/US99/30911
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US99/30999
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US00/00219
 PRIOR FILING DATE: 2000-01-05
 NUMBER OF SEQ ID NOS: 423
 SEQ ID NO 164
 LENGTH: 476
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-905-291A-164

Alignment Scores:

Pred. No.: 2,92e-21 Length: 476
 Score: 360.50 Matches: 141

Mon Nov 24 13:41:03 2003

Percent Similarity:	39.22%	Conservative:	70
Best Local Similarity:	26.21%	Mismatches:	200
Query Match:	11.91%	Indels:	128
DB:	10	Gaps:	23

US-09-712-338-1 (1-1668) x US-09-905-291A-164 (1-476)

[illegible]

RESULT 7

US-09-902-853-164
: Sequence 164, Application US/09902853
: Publication No. US20020192659A1
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, A.
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth, J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Mather, Jennie P.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William, I.
: TITLE OF INVENTION: Secretd and Transmembrane Polypeptides and Nucleic


```

; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-164

Alignment Scores:
Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservative: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 10 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-907-824-164 (1-476)
QY 50 GGGCCCTTCAGGAAGTACACGGCGCTCGTGGTAGAAGACAGTACCCAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
QY 110 CCGGGGTCAAGACTCTTACAAACCGCAAAACAATGTCACCA-----TCGGTACAGG 160
Db 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle 54
QY 161 AACCCGGGCAGAGGGCGTCTCGAGACTACGAGTACCCCG-----GGTGCMAATCCTAC 210
Db 55 GlnLysGlyArgGlnLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
QY 211 TCTGGATAT-----GTCGACACTCTCCGAGTCCCATACCTTCTTCTGTTCTTCCAA 264
Db 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePro 94
QY 265 GCCAGACATAACCCAGAAACTGCACCTATCACATTTGTGGTGAATGGTGGCCCTGGAAGC 324
Db 95 AlaGlnIleGlnProGluAspAlaProValIleValLeuThrLeuGlnGlyProGlyGly 114
QY 325 GATTCTTTGATCGGTCTCTTCGAAAGAGTTGGGCGCTTGGCCATTCATTCGACTTTGAT 384
Db 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr 134
QY 385 GACTACATCAACCTCTACTCGTGAACAGAGTCTCCAAATTTACTATTCTCTCCAGCCA 444
Db 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
QY 445 TTGGGAGTCGGCTTTTCATATAGTGATACGGTTGATGGTCCATTAACCCCTGTAACTGGG 504
Db 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly----- 167
QY 505 GTCGTGAAATTCGAGCTTTGAGAGGTTTGCAGGGCGGTACCCCAACCAATGATGCCACT 564
Db 168 -----Tyr 168
QY 565 CTGATCGATACTACCAATCTTTCGCGCAGAGCGCGCTTGGGAGATCTCTCAAGGATTCCTT 624
Db 169 AlaValAsnGluAspAspValAlaAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe 187
QY 625 AGTGGACTACCTAGCTTGGACTCTTAGGGTGCAGTCTAAGGACTTCAGTCTATGGACGGAG 684
```

```

; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-164

RESULT 8
US-09-907-824-164
; Sequence 164, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertrits, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tomas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
```

RESULT 9

```

QY      50 GGGCCCTTCAGGAAGTACACCGCGGTCGTCGTAGAAGACGACTACCCACAAGCCCCA   109
Db      22 GlyLeuPheArgSerLeuTyArgSerValSerMet-----Pro 34
               |||::: |||::: |||:::
QY     110 CCGGGGTCAAGACTCTTACACGCCCAACAATGTCACCA-----TCGGTGACAAGG   160
               |||    |||    |||    |||    |||    |||

```

```

1  APPLICANT: Ashkenazi, Avi
2  APPLICANT: Botstein, David
3  APPLICANT: Desnoyers, Luc
4  APPLICANT: Eaton, Dan L.
5  APPLICANT: Ferrara, Napoleone
6  APPLICANT: Filvaroff, Ellen
7  APPLICANT: Fong, Sherman
8  APPLICANT: Gao, Wei-Qiang
9  APPLICANT: Gerber, Hanspeter
10 APPLICANT: Gerritsen, Mary E.
11 APPLICANT: Goddard, A.
12 APPLICANT: Godowski, Paul J.
13 APPLICANT: Grimaldi, Christopher J.
14 APPLICANT: Gurney, Austin L.
15 APPLICANT: Hillan, Kenneth, J.
16 APPLICANT: Kljavin, Ivar J.
17 APPLICANT: Mather, Jennie P.
18 APPLICANT: Pan, James
19 APPLICANT: Paoni, Nicholas F.
20 APPLICANT: Roy, Margaret Ann.
21 APPLICANT: Stewart, Timothy A.
22 APPLICANT: Tumas, Daniel
23 APPLICANT: Williams, P. Mickey
24 APPLICANT: Wood, William, I.
25
26 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
27
28 FILE OF INVENTION: Acids Encoding the Same
29
30 FILE REFERENCE: 10466-14
31
32 CURRENT APPLICATION NUMBER: US/09/907,841
33
34 CURRENT FILING DATE: 2001-11-20
35
36 PRIOR APPLICATION NUMBER: PCT/US00/04414
37
38 PRIOR FILING DATE: 2000-02-22
39
40 PRIOR APPLICATION NUMBER: US 60/143,048
41
42 PRIOR FILING DATE: 1999-07-07
43
44 PRIOR APPLICATION NUMBER: US 60/145,698
45
46 PRIOR FILING DATE: 1999-07-26
47
48 PRIOR APPLICATION NUMBER: US 60/146,222
49
50 PRIOR FILING DATE: 1999-07-28
51
52 PRIOR APPLICATION NUMBER: PCT/US99/20594
53
54 PRIOR FILING DATE: 1999-09-08
55
56 PRIOR APPLICATION NUMBER: PCT/US99/20944
57
58 PRIOR FILING DATE: 1999-09-13
59
60 PRIOR APPLICATION NUMBER: PCT/US99/21090
61
62 PRIOR FILING DATE: 1999-09-15
63
64 PRIOR APPLICATION NUMBER: PCT/US99/21547
65
66 PRIOR FILING DATE: 1999-09-15
67
68 PRIOR APPLICATION NUMBER: PCT/US99/23089
69
70 PRIOR FILING DATE: 1999-10-05
71
72 PRIOR APPLICATION NUMBER: PCT/US99/28214
73
74 PRIOR FILING DATE: 1999-11-29
75
76 Remaining Prior Application data removed - See File Wrapper or PALM.
77
78 NUMBER OF SEQ ID NOS: 423
79
80 SEQ ID NO 164
81
82 LENGTH: 476
83
84 TYPE: PRT
85
86 ORGANISM: Homo sapiens
87
88 US-09-907-841-164

```

Alignment Scores:		
Pred. No.:	2.92e-21	Length:
Score:	360.50	Matches:
Percent Similarity:	39.22%	Conservative:
Best Local Similarity:	26.21%	Mismatches:
Query Match:	11.91%	Indels:
DB:	10	Gaps:

US-09-712-338-1 (1-1668) x US-09-907-841-164 (1-476)

Mon Nov 24 13:41:03 2003

35	ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle	54
161	AACCGGGGAGAGGGCGTCTCGAGACTACCCCG-----GGTGTCAAATCCTAC	210
55	GlnLysGlyArgGluLeuValGlyProPheProGlyLeuAsnMetLysSerTyr	74
211	TCTGATAT-----GTCCACACCTCTCCCGAGTCCCATCTTCTTCTGGTCTTCGAA	264
75	AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPhePheTrpPhePro	94
265	GCCAGACATACCCAGAACTGCACCTATCACATGTGGTGAATGTGGCCCTGGAGC	324
95	AlaGlnIleGlnProGluAspAlaProValValLeuTrpLeuGlnGlyGlyProGlyGly	114
325	GATTCTTTGATCGGTCTCTTGAAGAGTTGGCCCTTCCCATGTCAATTCCACTTTTGAT	384
115	SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr	134
385	GACTACATCAACCCCTCACTCGTGACAGAGGTCTCCAATTTACTATTCTGTGCCAGCCA	444
135	LeuArgAspArgAspPheProTrpThrThrLeuSerMetLeuTyrIleAspAsnPro	154
445	TTGGAGAGTCGGCTTTTCATATAGTCATACGGTGTATGGTCCATTAACCTGTTACTGGG	504
155	ValGlyThrGlyPheSerPheThrAspAspThrHisGly-----	167
505	GTCTCGAAATTCGAGCTTTTCAGGAGTTTCAGGGCCGGTACCCCAACCATGATGCCACT	564
168	-----Tyr	168
565	CTGATCCATACTACCAATCTTCCGCGAGAGCGCTTGGGAGATCTCGAAGATTCTCT	624
169	AlaValAsnGluAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe	187
625	AGTGGACTACCTAGCTTGGACTTAGGTGCAGTCTAAGCACTTCAGTCTATGGACGGAG	684
188	-----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu	203
685	AGCTATGGGGCACTATGTCCTCGCATCTTCAATCAITTTTACGACGAGATGAGAGA	744
204	SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn	221
745	ATTGCCAACGGTAGTGAATGGTGTTCAGCTTAATTTCAACTCTCTGGGAATTATTAA	804
222	-----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp	237
805	GGCATCATCAGACAGGGGATCCAGGCCCTTACTACCTCAATTCGCTGGAACATAACC	864
238	GlyTyrSerAspProGluSerIleIleGlyGlyTyrAlaGluPhe-----Leu	253
865	TACGGTATCAAGGTGTCAACAGAGACCGTCTCAACTACATGAAGTTTGCACCAAAATG	924
254	TyrGlnIleGlyLeuLeuAspGluLysGlnLysLysTyr-----PheGlnLysGln	270
925	CCAAATGGTGTCCAGGATTTGATTTCCACTCCACTGCCAAACAGACAAAC-----	975
271	-----CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe	287
976	GCATTAGTACTAGCCCTCTCGCCGCGAAGCCACCAC-----ATGTCAGGGACAAT	1029
288	GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn	307
1030	GTTCAGGG-----CCATACTACGCCCTTTGCTGGTGGTGTGTATGATATTCGGCAT	1083
308	ValThrGlyCysSerAsnTyrTyrAsnPheLeu-----	318
1084	CCATATGATGACCCGACTCCGCCAAGTTATTACAAACAAATTTCTGGCAAGGACTCTGTC	1143
319	ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal	338
1144	ATGGACGCTATCGGGGTCAACACTCAACTACACCCAGTCCCAATATGACTCTAC-----	1197
339	ArgGlnAlaIleHisValGly-----AsnGlnThrPheAsnAspGlyThrIleValGluLys	357

RESULT 10

US-09-904-011-164

Sequence 164, Application US/09904011
Publication No. US20030003530A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary B.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/904,011

CURRENT FILING DATE: 2001-07-11

PRIOR APPLICATION NUMBER: 09/665,350

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: US 60/146,222

:	PRIOR FILING DATE:	1999-07-28
:	PRIOR APPLICATION NUMBER:	PCT/US99/20594
:	PRIOR FILING DATE:	1999-09-08
:	PRIOR APPLICATION NUMBER:	PCT/US99/20944
:	PRIOR FILING DATE:	1999-09-13
:	PRIOR APPLICATION NUMBER:	PCT/US99/21090
:	PRIOR FILING DATE:	1999-09-15
:	PRIOR APPLICATION NUMBER:	PCT/US99/21547
:	PRIOR FILING DATE:	1999-09-15
:	PRIOR APPLICATION NUMBER:	PCT/US99/23089
:	PRIOR FILING DATE:	1999-10-05
:	PRIOR APPLICATION NUMBER:	PCT/US99/28214
:	PRIOR FILING DATE:	1999-11-29
:	PRIOR APPLICATION NUMBER:	PCT/US99/28313
:	PRIOR FILING DATE:	1999-11-30
:	PRIOR APPLICATION NUMBER:	PCT/US99/28564
:	PRIOR FILING DATE:	1999-12-02
:	PRIOR APPLICATION NUMBER:	PCT/US99/28565
:	PRIOR FILING DATE:	1999-12-02
:	PRIOR APPLICATION NUMBER:	PCT/US99/30095
:	PRIOR FILING DATE:	1999-12-16
:	PRIOR APPLICATION NUMBER:	PCT/US99/30911
:	PRIOR FILING DATE:	1999-12-20
:	PRIOR APPLICATION NUMBER:	PCT/US99/30999
:	PRIOR FILING DATE:	1999-12-20
:	PRIOR APPLICATION NUMBER:	PCT/US00/00219
:	PRIOR FILING DATE:	2000-01-05
:	NUMBER OF SEQ ID NOS:	423
:	SEQ ID NO	164
:	LENGTH:	476
:	TYPE:	PRT
:	ORGANISM:	Homo Sapien
:	US-09-904-011-164	

Alignment Scores:
 Pred. No.: 2, 92e-21 Length: 476
 Score: 360.50 Matches: 141
 Percent Similarity: 39..22% Conservative: 70
 Best local Similarity: 26..21% Mismatches: 200
 Query Match: 11.91% Indels: 128
 DB: 11 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-904-011-164 (1-476)

QY	50	GGGCCCTTCCAGGAATACACCGGGCTCCGTCCGTGAAGAAGACAGCTACCACAAGACCCC	109
Db	22	GlyLeuPheArgSerLeuTyrrArgSerValSerMet-----Pro	34
QY	110	CGGGGTCAAGACTCTTACAACCGCAACAATGTACCA-----TCCGTACAAGG	160
Db	35	ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrrIleGluAlaGlyLysIle	54
QY	161	AACCGGGCAGAGGGCTTCGAGACTACCCCG-----GGTGTCAAATCCTTAC	210
Db	55	GlnLysGlyArgGluSeuSerLeuValGlyProPheProGlyLeuAsnMetLysSeryr	74
QY	211	TCTGGATAT-----GTCGACACTCTCCGAGTCCCCACTCTCTCTGGTTCTTCGAA	264
Db	75	AlaGlyPheLeuThrValAsnLysThrTyrrAsnSerAsnLeuPheHeirpPhePro	94
QY	265	GCCAGACATAACCCAGAAATGCACCTATCATATTGTGGTGAATGGTGGCCCTCGAAGC	324
Db	95	AlaGlnIleGlnProGluAspAlaProValLeuTipLeuGlnGlyGlyProGlyGly	114
QY	325	GATTCTTTGATCGGTCTCTTCGAGAGTGGGCCCTTCGCATGTCAATTTCGACTTTTGAT	384
Db	115	SerSerMetPheGlyLeuPheValGluHisGlyProTyrrValValThr-SerAsnMetThr	134
QY	385	GACTACATCAACCCCTCACTCCGTGGAACGAGGTCTCCAATTTTACTATTCTGTCCAGCCA	444
Db	135	LeuArgAspArgAspPheProTipThrrThrrLeuSerMetLeuTyrrIleAspAsnPro	154
QY	445	TTGGGAGTCGGCTTTTCATATAGTACATACGGTTGATGGTCCATTAAACCCGTGAAC	504

Db 435 AlaGlyAspPheHisGlnValIleLeuArgGlyGlyHisIleLeuProTyrAspGln 454
Qy 1495 CCATCGCTCCCTGCATTTAAACGGGACTATCTTC-----GGTTGGGAT 1542
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp 472

RESULT 11

US-09-906-742-164
; Sequence 164, Application US/09906742
; Publication No. US20030023054A1

GENERAL INFORMATION:

- APPLICANT: Genentech, Inc.
- APPLICANT: Ashkenazi, Avi
- APPLICANT: Botstein, David
- APPLICANT: Desnoyers, Luc
- APPLICANT: Eaton, Dan L.
- APPLICANT: Ferrara, Napoleone
- APPLICANT: Filvaroff, Ellen
- APPLICANT: Fong, Sherman
- APPLICANT: Gao, Wei-Qiang
- APPLICANT: Gerber, Hanspeter
- APPLICANT: Gerritsen, Mary E.
- APPLICANT: Goddard, A.
- APPLICANT: Godowski, Paul J.
- APPLICANT: Grimaldi, Christopher J.
- APPLICANT: Gurney, Austin L.
- APPLICANT: Hillan, Kenneth, J.
- APPLICANT: Kiljavin, Ivar J.
- APPLICANT: Mather, Jennie P.
- APPLICANT: Pan, James
- APPLICANT: Paoni, Nicholas F.
- APPLICANT: Roy, Margaret Ann
- APPLICANT: Stewart, Timothy A.
- APPLICANT: Tumas, Daniel
- APPLICANT: Williams, P. Mickey
- APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

Acids Encoding the Same

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/906,742

PRIOR FILING DATE: 2001-07-16

PRIOR APPLICATION NUMBER: 09/665,350

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: US 60/146,222

PRIOR FILING DATE: 1999-07-28

PRIOR APPLICATION NUMBER: PCT/US99/20594

PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-13

PRIOR APPLICATION NUMBER: PCT/US99/21090

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/23089

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: PCT/US99/28214

PRIOR FILING DATE: 1999-11-29

PRIOR APPLICATION NUMBER: PCT/US99/28313

PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: PCT/US99/28564

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/28565

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/30095

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: PCT/US99/30911

PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-164

Alignment Scores:

Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservative: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 11 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-906-742-164 (1-476)

Qy	50	GGGCCCCCTCCAGGAAGTACACGGCGTCGTCGGTAGAGACGCTACCCAGAACCCCA	109
Db	22	GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro	34
Qy	110	CGGGGTCAGACTCTTACACCGCAACAATGTCAACA-----TCCGGTACAAGG	160
Db	35	ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle	54
Qy	161	AACCCGGGCGAGAGGGCGTCGCGAGACTACCCCG-----GGTGTCAATCTCTAC	210
Db	55	GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr	74
Qy	211	TCTGGATAT-----GTCGACACCTCTCCGAGTCCCATACCTCTCTCTGGTCTTCGAA	264
Db	75	AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePhePro	94
Qy	265	GCAGACATAACCCAGAACTGCACCTATCACATCTGTGGTTGAATGGTGGCCCTGGAAGC	324
Db	95	AlaGlnIleGlnProGluAspAlaProValValLeuThrLeuGlnGlyGlyProGlyGly	114
Qy	325	GATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTGGCCATGTCATTCGATTTTGAT	384
Db	115	SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr	134
Qy	385	GACTACATCAACCCCTCACTCGTGGACGAGGTCCCAATTACTATTCTCTGCCACCCA	444
Db	135	LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro	154
Qy	445	TTGGGACTCGGCTTTTCATATAGTACGGTTGATGGGTCCATTAAACCCCTGTAACCTGGG	504
Db	155	ValGlyThrGlyPheSerPheThrAspAspThrHisGly-----	167
Qy	505	GTGCTCGAAAATTCGAGCTTTGCAGAGATTTCAGGGCCGGTACCCCAACATTGTGCCACT	564
Db	168	-----Tyr	168
Qy	565	CTGATCGATATACCAATCTTCCGCGACAGCCCGCTTGGGAGATCTCGCAGGATTCCTT	624
Db	169	AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe---	187
Qy	625	AGTGGACTACCTAGCTTGGACTCTAGGGTGCAGTCTAAGGACTTCAGTCTATGACCGAG	684
Db	188	-----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu	203
Qy	685	AGCTATGGAGGCGACTATGGTCTCGATCTTCAATCATTTTACGACGACAAATCAGAGA	744
Db	204	SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn-----	221
Qy	745	ATTGCCAACGGTAGTGTAAATGGTCTTCAGCTTAATTTCACTCTCTGGGAATTATTAAAC	804
Db	222	-----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp	237

Qy 50 GGGCCCTCCAGGAAGTACACCGCGTCGGTAGAAGACAGCTACCCAGAACCCCA 109

QY	1144	ATGGAGCGTATCGCGTCACATCAACTACACCGCTCCAAATAATGACGCTTAC-----	1197
Db	339	ArgGlnAlaIleHisValGly---AsnGlnThrPheAsnAspGlyThrIleValGluLys	357
QY	1198	TAGCGTTTCCAGCAACAGCGGACATTGTCTGCTGCCCAACTTCATCGAAGACCTCGAGGAG	1257
Db	358	TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn	377
QY	1258	ATCCTTGCTCTCCCGTGGTGTCTCCCTCATTATGGCGACCGCATTCATC-----	1311
Db	378	-----TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla	392
QY	1312	-----TGCAACTGGTTCCGCGCTCAGCGCGTTTCC	1341
Db	393	AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTyrLysGlySerGln-----	409
QY	1342	CTCGCTCGCAACTACTCCCAAGCGCGCCAG-----TTCCGAAGC-----	1380
Db	410	-----GluTyrLysLysAlaGluLysValTyrLysIlePheLysSerAspSer	426
QY	1381	-----GCAGGGTACAGCCCTCGAAAGTCAACGGCGCTCGAGTATGGGAACTCGCGAG	1434
Db	427	GluValAlaGlyTyr-----IleArgGln	434
QY	1435	TATGGTAAATTTCCCTTCATCTCGCGTCTATGAGCAGCGCATGAAGTCCCATCTACCAG	1494
Db	435	AlaGlyAspPheHisGlnValIleIleArgGlyGlyGlyHisIleLeuProTyrAspGln	454
QY	1495	CCATCGCCCTCCCTGCAATTGTTTAAACGGACATCTTTC-----GGTTGGGAT	1542
Db	455	ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp	472

RESULT 13

US-09-907-613-164

Sequence 164, Application US/09907613

Publication No. US20030027145A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi

APPLICANT: Botstein, David

APPLICANT: Desnovers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, A.

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth, J.

APPLICANT: Klijavin, Ivar J.

APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/907,613

CURRENT FILING DATE: 2001-07-17

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

QY	445	TTGGAGTCGGCTTTTCATATAGTCATACGGTTGATGGCTCCATTAAACCTGTAACTGGG	504
Db	155	ValGlyThrGlyPheSerPheThrAspThrHisGly	167
QY	505	GTCGTCGAAATTCGAGCTTTGCAGAGTTTCAGGCCGGTACCGAACCAATTGATGCCACT	564
Db	168	-----Tyr	168
QY	565	CTGATCGATACTACCAATCTTTGCCGACAGGCGCTTGGGAGATCCTGCAAGATTCCCT	624
Db	169	AlaValAsnGluAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe	187
QY	625	AGTGGACTACCTAGCTTGGACTTAGGTCGAGCTCTAAGGACTTCAGTCTATGGACGGAG	684
Db	188	-----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu	203
QY	685	AGCTATCGAGGCACTATGGTCTCGCATCTTCAATCATTTTACGAGCAGACATGAGAGA	744
Db	204	SerTyrAlaGlyTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn	221
QY	745	ATTGCCAACGCTAGTGTAAATGGTTTCAGCTTAAATTCACACTCTCTGGGAATTATTAAC	804
Db	222	-----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp	237
QY	805	GGCATCATCGACGAGCGATCCAGGCCCTTACTACCTGAATTCGCTGTGAACAATACC	864
Db	238	GlyTyrSerAspProGluSerIleIleGlyGlyTyrAlaGluPhe	253
QY	865	TACGGTATCAAGCTGTCAAGACACCGTCTACAACTACATGAAGTTGCCAACCAATG	924
Db	254	TyrGlnIleGlyLeuLeuAspGluLysGlnLysTyr	270
QY	925	CCAAATGGTTCGACGAGTTGATTTCCACCTCGCAACAGACAAC	975
Db	271	-----CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe	287
QY	976	GCATTAGCTGACTACGCCCTCTCGCGCAAGCACCACAAAC	1024
Db	288	GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn	307
QY	1030	GTGAGGGG	1083
Db	308	ValThrGlyCysSerAsnTyrTyrAsnPheLeu	318
QY	1084	CCATATGATGACCCGACTCGGCCAAGTTATTACACAAATTTCTGGCAAGGACTCTGTC	1143
Db	319	ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal	338
QY	1144	ATGGACGTATCGGGTCAACATCACTACCCAGTCCAATATGACGCTAC	1197
Db	339	ArgGlnAlaIleHisValGly	357
QY	1198	TACGCTTTCCAGCAACAGGCGACTTTGTCTGGCCCACTTCATCGAAGACCTCGAGGAG	1257
Db	358	TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn	377
QY	1258	ATCCTTGCTCTCCCGTGGTCTCCCTCATCTACTACCGCGCCGATCATATC	1311
Db	378	-----TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla	392
QY	1312	-----TGCAACTGGTTCGGCGGTGAGCGCGTTTCC	1341
Db	393	AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTrpLysGlySerGln	409
QY	1342	CTCGTGGCACTACTCCCAAGCCGCCAG	1380
Db	410	-----GluTyrLysLysAlaGluLysLysValTrpLysIlePheLysSerAspSer	426
QY	1381	-----GCAGGGTACACGCCCTGAAGTCAACGGCGTTCGATATGGGGAACCTCGGAG	1434
Db	427	GluValAlaGlyTyr	434

QY 1435 TATGTAATTTCTCTCACTCGCTCTATGAGCAGCCGATGAGTCCCATATACAG 1494
Db 435 AlaglyAspPheHisGlnValleileArgGlyGlyHisleLeuProTyrAspGln 454
QY 1495 CCCATCGCTCCCTGCAATTTTAAACCGGACTATCTC-----GGTTGGAT 1542
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp 472

RESULT 14
US-09-907-942-164
; Sequence 164, Application US/09907942
; Publication No. US20030027146A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mackey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,942
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-942-164
Alignment Scores: 2.92e-21 Length: 476
Pred. No.: 360.50 Matches: 141
Score: 39.22% Conservative: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 11 Gaps: 23
US-09-712-338-1 (1-1668) x US-09-907-942-164 (1-476)
QY 50 GGGCCCTTCCAGGAAGTACACCGGCGTCGCTGAGACAGCTACCCAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
QY 110 CCGGGGTCAAGACTCTTACACCGCAACAATGTCCACCA-----TCCGGTACAAGG 160
Db 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle 54
QY 161 AACCCGGGCGAGAGGGCGTCTCGAGACTACCCCG-----GGTCAAAATCCTAC 210
Db 55 GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
QY 211 TCTGGATAT-----GTCACACCTCTCCGAGTCCCATACCTTCTTCTGTTCTTCGAA 264
Db 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePro 94
QY 265 GCAGACATAACCCAGAACTGCACCTATCACATTTGTGTGAATGGTGGCCCTCGAAGC 324
Db 95 AlaGlnIleGlnProGluAspAlaProValValLeuThrLeuGlnGlyProGlyGly 114
QY 325 GATTCTTTGATCGGTCTCTTCGAGAGTGGGCGCTTCGATGTCATGTCATTCGACTTTGAT 384
Db 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyrValThrSerAsnMetThr 134
QY 385 GACTACATCAACCTCACTCTCGAAGAGGTCTCCAAATTACTATTCTGTGCCAGCCA 444
Db 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
QY 445 TTGGGAGTCGGCTTTTCATATAGTATACGGTTGATGGGTCCATTAAACCTGTAACTGGG 504
Db 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly-----167
QY 505 GTCGTCGAAAATTGAGCTTTTGACAGAGTTTCAGGGCGGTACCCCAACCATTTGTCGACT 564
Db 168 -----Tyr 168
QY 565 CTGATCGATACCTACCAATCTTCCGCGAGAGCGCGCTTGGGAGATCCTGCAAGGATTCCTT 624
Db 169 AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe--- 187
QY 625 AGTGGACTACCTAGCTTGGACTCTAGGGTGCAGGTCTAAGGACTTCAGTCTATGACCGAG 684
Db 188 -----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu 203
QY 685 AGCTATGAGGCGCACTATGCTCTGCAATCTTCAATCATTTTACGACGACGAATGAGAGA 744
Db 204 SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn----- 221
QY 745 ATTGCCAACGGTAGTGTATGTTGTTGAGTCTTCACTTATTTCACTCTCTGGGATATTAAAC 804
Db 222 -----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp 237

QY 805 GGCATCATCGACGAGCGGTACTACCCTGGTAATTCCCTGTGAACAATACC 864
 |||
Db 238 GlyTyrSerAspProGluSerIleileGlyGlyTyraAlaGluphe-----Leu 253
 |||
QY 865 TAGCGTTCAACGGCTGTCAACAGACCGTCTAACAACTCATGAAGTTTGCCAACCAAATG 924
 |||
Db 254 TyrGlnIleGlyLeuLeuAaspGluLysGlnLysLysTyr-----PheGlnLysGln--- 270
 |||
QY 925 CCAAAATGGTTCGCCAGATTTGATTTCCACTGCCAACAGACAAC-----CGCACC 975
 |||||
Db 271 -----CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe 287
 |||
QY 976 GCATTAGTGTACTACGCCCTCTGCGCGAAGCACCAAC-----AUTGTCAGGGACAAT 1029
 :: |||
Db 288 GluIleLeuAaspLysLeuLeuAaspGlyAspLeuthrSerAspProSerTyrPheGluAsn 307
 |||
QY 1030 GTTAGAGGG-----CCATACTACGCCCTTGCTGCTGTGTATGATATTCCGCAT 1083
 |||
Db 308 ValThrGlyCysSerAsnTyrTyrAsnPheLeu----- 318
 |||
QY 1084 CCATATGATGACCCCGACTCCGCCAAGTTATTACAAACAAATTCCTGGCAAAGGACTCTGTC 1143
 :::|||
Db 319 ArgCysThrGluProGluAaspGlnLeuTyrValLysPheLeuSerLeuProGluVal 338
 |||
QY 1144 ATGACCGCTATCGCGTCAACATCAACTACACCCAGTCCAATATACGCTAC----- 1197
 :: |
Db 339 ArgGlnAlaIleHisValGly--AsnGlnThrPheAsnAspGlyThriIeValGluLys 357
 |||
QY 1198 TACGCTTTCAGCAACAGGCGACTTGTCTGCGCCCAACTTCATCGAAGACCTCGAGGAG 1257
 :: |||
Db 358 TyrLeuArgGluAspThrValGlnSerValLysProTrpLeuThrGluIleMetAsnAsn 377'
 |||
QY 1258 ATCCTTGCTCTCCC CGCGTCTCCCTCATCTATGGCGACCGCATTCATC- 1311
 :: |||
Db 378 -----TyrLysValLeuIleTyrAsnGlyLsnLeuAaspIlelleValala 392
 |||
QY 1312 -----TGCAACTGTTTCGCGGTCTAGGCGGTTCCTCC 1341
 |||
Db 393 AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTrpLysGlySerGln----- 409
 |||
QY 1342 CTCGCTCGGAECTACTCCCAAGCGGCCAG-----TTCGGAAGC----- 1380
 |||
Db 410 -----GluTyrLysLysAlaGluLysLysValTrpLysIlePheLysSerAspSer 426
 |||
QY 1381 ----CGAGGTACACGCCCTCAAAGTCAACGGCGTCAAGTATCGGAACTCGCGAG 1434
 |||
Db 427 GluValAlaGlyTyr-----lleArgGln 434
 |||
QY 1435 TATGGTAATTCTCTCTTCATCTCGCGTCTATGAGGCGGCGCATCAAGTCCCATCTACGAG 1494
 |||
Db 435 AlaGlyAspPheHisGlnValIlelleArgGlyGlyHisIleleuProTyrAspGln 454
 |||
QY 1495 CCCATCGCTCCCTGCANTGTTTAACCGGACTATCTTC-----GGTTGGAT 1542
 |||
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTrpAsp 472

RESULT 15

US-09-796-753-40
; Sequence 40, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246

Search completed: November 21, 2003, 17:49:19
Job time : 85.5 secs

QY	211	TCTGATAT-----GTGCACACTCTCCGAGTCCCACATCCTCTTCCTGGTTCTTGAA	264
Db		::: ::: ::: ::: ::: ::: ::: :::	
QY	75	AlaGlyPheLeuThrValAsnLysThrTyAsnSerAsnLeuPhePheTrpPhePro	94
QY	265	GCCAGACATAACCAGAATCGCACATATCATATTGGTTGAATGGTGGCCCTGGAAGC	324
Db		::: ::: ::: ::: ::: ::: :::	
QY	95	AlaGlnIleGlnProGluAspAlaProValValLeuTrpLeuGlnGlyGlyProGlyGly	114
QY	325	GATCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTGCCATGTCAATTGCACTTTTGAT	384
Db		::: ::: ::: ::: ::: ::: :::	
QY	115	SerSerMetPheGlyLeuPheValGluHisGlyProTyValValThrSerAsnMetThr	134
QY	385	GACTACATCAACCCCTCACTCTGTGGAACGAGGTCTCCAATTACTATCTGTGCCGCGA	444
Db		::: ::: ::: ::: ::: ::: :::	
QY	135	LeuArgAspArgAspPheProTrpThrThrLeuSerMetLeuTyrlleAspAsnPro	154
QY	445	TTGGAGTCCGGCTTTTCATATAGTAGTATACGGTTGATGGGTCCATTAAACCTGTAACTGGG	504
Db		::: ::: ::: ::: ::: ::: :::	
QY	155	ValGlyThrGlyPheSerPheThrAspAspThrHisGly-----	167
QY	505	GTCTGCGAAAAATTCGAGCTTTGACGAGTTTCAGGCGGTACCGACCATGTATGCCACT	564
Db		::: ::: ::: ::: ::: ::: :::	
QY	565	CTGATCGATACCAATCTTGGCGCAGAGCGCTTGGGAGATCTCTGCAAGGATTCCTT	624
Db		::: ::: ::: ::: ::: ::: :::	
QY	169	AlaValAsnGluAspAspValAlaArgAspLeuTySerAlaLeuIleGlnPhePhe---	187
QY	625	AGTGGACTACCTAGCTTGGACTCTAGGGTCGACGTAAAGCACTTCAGTCTATGACGAG	684
Db		::: ::: ::: ::: ::: ::: :::	
QY	188	-----GlnIlePheProGluTyLysAsnAsnAspPheTyValThrGlyGlu	203
QY	685	AGTATGGAGGCACTATGTCTCGTCAATCTTCAATCATTTTACGAGCAAGTAAGAGA	744
Db		::: ::: ::: ::: ::: ::: :::	
QY	204	SerTyAlaGlyLysTyValProAlaIleAlaHisLeuIleHisSerLeuAsn----	221
QY	745	ATTGCCAACGGTAGTGTATGTTGTGTTCAGCTTAATTTCAACTCTCTGGGAATTTAAAC	804
Db		::: ::: ::: ::: ::: ::: :::	
QY	222	-----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp	237
QY	805	GGCATATCGACGAGGCGATCCAGGCCCTTACTACCCCTGAATTCGTGTGCAACAATACC	864
Db		::: ::: ::: ::: ::: ::: :::	
QY	238	GlyTySerAspProGluSerIleIleGlyTyAlaGluPhe-----Leu	253
QY	865	TACGGTATCAAGCTGTCAACGAGACCGCTACAACTACATGAAGTTTGCACCAACAAATG	924
Db		::: ::: ::: ::: ::: ::: :::	
QY	254	TyrGlnIleGlyLeuLeuAspGluLysGlnLysTySyr-----PheGlnLysGln---	270
QY	925	CCAAATGGTTCGCCAGGATTTGATTTCCACTTCGCAACACAGACAAAAC-----CGCACC	975
Db		::: ::: ::: ::: ::: ::: :::	
QY	271	-----CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe	287
QY	976	GCATTAGTGACTAGCCCTCTGCGCCGAGGCCACCAAC-----ATGTGCGAGGACAAAT	1029
Db		::: ::: ::: ::: ::: ::: :::	
QY	288	GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspSerTyPheGlnAsn	307
QY	1030	GTTCAGGGG-----CCATATCTACGCCCTTTGCTGGTGGTGTATGATATTCGGCAT	1083
Db		::: ::: ::: ::: ::: ::: :::	
QY	308	ValThrGlyCysSerAsnTyTyAsnPheLeu-----	318
QY	1084	CCATATGATGACCCGACTCCGCGCAAGTTATTACAAATTTCTGCRAAGGACTCTGTC	1143
Db		::: ::: ::: ::: ::: ::: :::	
QY	319	ArgCysThrGluProGluAspGlnLeuTyTyValLysPheLeuSerLeuProGluVal	338
QY	1144	ATGACGCTATCGGGCTCAACATCACTACACCCAGTCCCAATATATGACCTCTAC	1197
Db		::: ::: ::: ::: ::: ::: :::	
QY	339	ArgGlnAlaIleHisValGly---AsnGlnThrPheAsnAspGlyThrIleValGluLys	357
QY	1198	TACGTTCTCCAGAAACAGCGGACTTTGTCTGGCCCACTTCATCGAAGCACTCGAGGAG	1257
Db		::: ::: ::: ::: ::: ::: :::	
QY	358	TyrLeuArgGluAspThrValGlnSerValLysProTrpLeuThrGluIleMetAsnAsn	377
Db		::: ::: ::: ::: ::: ::: :::	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:37:34 ; Search time 29 Seconds
(without alignments)

Title: US-09-712-338-1
Perfect score: 3027
Sequence: 1 atcgctggctacgaattctt.....ccagtggtggtatgcatag 1668

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model
-Q=/cgn2_1/USPTO.spool_p/US09712338/tunat 17112003 170149 10565/app_query.fasta_1.1863
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CCN 1.1.31 @runat 17112003 170149 10565 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:**
1: /cgn2_6/prodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/prodata/2/iaa/6C.COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2975	98.3	554	3	US-08-943-714-2
2	894	25.5	423	3	US-08-943-714-9
3	630	20.8	557	1	US-08-309-341-2
4	630	20.8	557	1	US-08-608-267-2
5	630	20.8	557	1	US-08-608-452-2
6	630	20.8	557	1	US-08-608-224-2
7	630	20.8	557	2	US-08-967-149-2
8	623	20.6	557	1	US-08-309-341-4
9	623	20.6	557	1	US-08-608-267-4
10	623	20.6	557	1	US-08-608-452-4
11	623	20.6	557	1	US-08-608-224-4
12	623	20.6	557	2	US-08-967-149-4

13	568.5	18.8	491	1	US-09-640-305-4	Sequence 4, Appli
14	568.5	18.8	491	1	US-08-360-673-4	Sequence 4, Appli
15	523.5	17.3	532	3	US-08-899-324-33	Sequence 33, Appl
16	523.5	17.3	532	3	US-08-329-892B-33	Sequence 33, Appl
17	513.5	17.0	421	2	US-08-807-263-4	Sequence 4, Appli
18	361.5	11.9	476	2	US-08-828-488-3	Sequence 3, Appli
19	361.5	11.9	476	4	US-09-299-689A-3	Sequence 3, Appli
20	349	11.5	477	2	US-08-828-488-1	Sequence 1, Appli
21	349	11.5	477	4	US-09-299-689A-1	Sequence 1, Appli
22	337	11.1	480	2	US-08-828-488-8	Sequence 8, Appli
23	337	11.1	480	4	US-09-299-689A-8	Sequence 8, Appli
24	337	11.1	480	4	US-09-702-705-336	Sequence 336, App
25	337	11.1	480	4	US-08-828-488-7	Sequence 336, App
26	312.5	10.3	471	2	US-08-828-488-7	Sequence 7, Appli
27	312.5	10.3	471	4	US-09-299-689A-7	Sequence 7, Appli
28	267	8.8	523	3	US-08-943-714-11	Sequence 11, Appli
29	246.5	8.1	446	1	US-08-665-966-10	Sequence 10, Appl
30	246.5	8.1	446	3	US-09-041-780-10	Sequence 10, Appl
31	208	6.9	530	3	US-08-943-714-12	Sequence 12, Appl
32	204.5	6.8	481	3	US-08-943-714-10	Sequence 10, Appl
33	193	6.4	35	3	US-08-943-714-5	Sequence 5, Appli
34	172	5.7	179	1	US-08-665-966-8	Sequence 8, Appli
35	172	5.7	179	3	US-09-041-780-8	Sequence 8, Appli
36	169.5	5.6	351	2	US-08-828-488-5	Sequence 5, Appli
37	169.5	5.6	351	4	US-09-299-689A-5	Sequence 5, Appli
38	163.5	5.4	878	4	US-09-556-706B-2	Sequence 2, Appli
39	161	5.3	907	3	US-08-783-774-2	Sequence 2, Appli
40	161	5.3	907	4	US-09-328-599A-1	Sequence 1, Appli
41	161	5.3	907	5	PCT-US95-04611A-19	Sequence 19, Appl
42	156	5.2	1048	4	US-09-171-699-10	Sequence 10, Appl
43	155.5	5.1	1911	4	US-09-854-856-64	Sequence 64, Appl
44	155.5	5.1	1939	4	US-09-854-856-48	Sequence 48, Appl
45	155.5	5.1	1971	4	US-09-854-856-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:

; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus

; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSER: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-943-714-2

Alignment Scores:
Pred. No.: 2,346-275 Length: 554
Score: 2975.00 Matches: 554
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.28% Indels: 0
DB: 3 Gaps: 0

US-09-712-338-1 (1-1668) x US-08-943-714-2 (1-554)

QY	1	AFPGCTGGCTACGAATTTCTCTCAGTGTACCCCTTGGTGGAGCCAGTTGGCCCTTCCA	60
DB	1	MetArgGlyTyrGluPheLeuSerValLeuProLeuValAlaAlaSerTrpAlaLeuPro	20
QY	61	GGAAGTACACCGGCTCGGTAGAACACAGTACCCACAGAACCCACCGGGGTCAAG	120
DB	21	GlySerThrProAlaSerValGlyArgArgGlnLeuProGlyAsnProThrGlyVallys	40
QY	121	ACTCTTACACCGCAACATTCACATCCGGTACAGGAACCCCGGGCAGAGGCGTC	180
DB	41	ThrLeuThrAlaAlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyVal	60
QY	181	TGGAGACATACCCCGGTGTCAAACTCTCTGTGATATGTGACACCTCTCCGAGTCC	240
DB	61	CysGluThrThrProGlyVallysSerTyrSerGlyTyrValAspThrSerProGluSer	80
QY	241	CATACCTTCTTGGTCTTCGAGCCAGACACATACCCAGAACTGCACCTATCACATTG	300
DB	81	HisThrPheThrPhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeu	100
QY	301	TGGTTCGAATGGTGGCTGGAAGCGATTCTTTCATCGGTCTCTTCAAGAGATTGGCCCT	360
DB	101	TrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluGlyPro	120
QY	361	TGCCATGTCAATTCGACTTTTGTATGACTACATCAACCTCAGTGGGACGAGGCTCC	420
DB	121	CysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSer	140
QY	421	AATTACTATCTCTGTCGACCATTCGGAGTGGCTTTTCATATAGTATGATACGTTGAT	480
DB	141	AsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAsp	160
QY	481	GGGTCCATTAAACCTGTAACCTGGGGTGTCTGAAATTCGAGCTTTGACGAGGTTCAAGGC	540
DB	161	GlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGly	180
QY	541	CGGTACCCAAACCATGTGATGCACTGTGATGATACCTACCTATCTTGGCGAGAGCGCT	600
DB	181	ArgTyrProThrIleAspAlaThrLeuIleAspThrAsnLeuAlaAlaGluAlaAla	200
QY	601	TGGAGATCTGTCAGAGGATTCCTAGTGGACTACCTAGCTTGGACTCTAGGCTGAGTCT	660
DB	201	TrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSer	220
QY	661	AAGGACTTCAGTCTATGGACGAGAGCTATGAGGGCACTATGCTCTGCAATTTCTCAAT	720
DB	221	LysAspPheSerLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPhePheAsn	240
QY	721	CATTTTACGACAGATGAGAGATTGCCACGGTACTGTTAATGGTGTTCAGCTTAAT	780

DB	241	HisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsn	260
QY	781	TTCAACTCTCTGGGAATTTAAGCGCATCATCGACGAGGATCCAGGCCCTTACTAC	840
DB	261	PheAsnSerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyr	280
QY	841	CCTGAATTCGTGTGAACAATACCTACGGTATCAAGCTGTCAACGAGACCGCTCTACAC	900
DB	281	ProGluPheAlaValAsnAsnThrTyrGlyIleIleAlaValAsnGluThrValTyrAsn	300
QY	901	TACATGAAGTTTGCACCAACCAATGCCAAATGGTTCGCCAGATTTGATTTCCACCTGCAA	960
DB	301	TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys	320
QY	961	CAGACAAACCCACCGCATTTAGCTGACTACCCCTCTCGCGGAGCCACCAACATGTC	1020
DB	321	GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys	340
QY	1021	AGGACAAATGTGAGGGGCCATACCTACGCCCTTTGCTGTGCTGTGTGTATGATTTCCG	1080
DB	341	ArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArg	360
QY	1081	CATCCATATGATGACCCGACTCCGCAAGTTATTACAAATTTCTGCAAGAGACCT	1140
DB	361	HisProTyrAspAspProThrProProSerTyrTyrAsnLysPheLeuAlaLysAspSer	380
QY	1141	GTTCATGAGCGTATCGGCTGCAACATCACTACACCCAGTCCCAATTAATGACCTCTACTAC	1200
DB	381	ValMetAspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyr	400
QY	1201	GCCTTCCAGCAACAGGCGACTTGTCTGGCCCAACTTCATCGAAGACCTCGAGGAGATC	1260
DB	401	AlaPheGlnGlnThrGlyAspPheValTyrProAsnPheIleGluAspLeuGluIle	420
QY	1261	CTTGCTCTCCCGTGGTGTCTCCCTCATCTATGCGGACGCGGATCTGCACTG	1320
DB	421	LeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrp	440
QY	1321	TTCGGGGTTCAGGCGGTTTCCCTCGCTGCGAATCTCCCAAGCCGCCAGTTCCGAAGC	1380
DB	441	PheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSer	460
QY	1381	GCAGGTTCACAGCCCTCGAAAGTCAACGGCGTCAAGTATGGGAAACTCCGAGTATGGT	1440
DB	461	AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly	480
QY	1441	AATTTCTCTCTACTCGGCTATGAGGCGGCGCATGAAGTCCCATACCTACGAGCCATC	1500
DB	481	AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle	500
QY	1501	GCCTCCCTGCAATTTGTTTAAACCGGACTATCTTCGGTGGGATATCGAGAGGCGCAAG	1560
DB	501	AlaSerLeuGlnLeuPheAsnArgThrIlePheGlyTrpAspIleAlaGluGlyGlnLys	520
QY	1561	AAGATCTGCGCCAGCTACAAAGACGAATGGAACCGCTACAGCTACGATACAGCTCGTCC	1620
DB	521	LysIleTrpProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer	540
QY	1621	GTGCGCTGCTACCGGTACCGATCTCCAGTCTCGGTATG	1662
DB	541	ValProLeuProThrAlaThrSerMetSerValGlyMet	554

RESULT 2
US-08-943-714-9
; Sequence 9, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik

; APPLICANT: Dambmann, Claus
 ; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 61875780 No. 61875780disk of No. 61875780th America, Inc.
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/943,714
 ; FILING DATE: 03-OCT-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lambiris, Elias J
 ; REGISTRATION NUMBER: 33,728
 ; REFERENCE/DOCKET NUMBER: 4990.200-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 423 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-943-714-9
 ;
 Alignment Scores:
 Pred. No.: 1,116-76 Length: 423
 Score: 894.00 Matches: 186
 Percent Similarity: 54.31% Conservative: 66
 Best Local Similarity: 40.09% Mismatches: 160
 Query Match: 29.53% Indels: 52
 DB: 3 Gaps: 8
 ;
 US-09-712-338-1 (1-1668) x US-08-943-714-9 (1-423)
 QY 175 GCGCTGTCGAGACTACCCGGGTGTCAAATCTCTACTCTGGATATCTGCACACCTCTCCC 234
 Db 6 GlylleCysGluThrThrProGlyValAsnGlnTyrSerGlyTyrLeuSerValGlySer 25
 QY 235 GAGTCCCATACCTTCTCTGTTCTTCCGAAGCCAGACATAACCCAGAACTGCACCTATC 294
 Db 26 AsnMetAsnMetTyrPheTyrPheGluAlaAsgAsnAsnProGlnGlnAlaProLeu 45
 QY 295 ACATTGTGTTCAATGGTGGCCCTGGAAGCGGATCTTTGATCGTCTCTTCGAAGAGTTG 354
 Db 46 AlaAlaTyrPheAsnGlyGlyProGlyCysSerSerMetIleGlyLeuPheGlnGluAsn 65
 QY 355 GGCCCTTCCCAT--GTCAATTTCGACTTTTGNATGACTACATCAACCTCATCTCGTGGAAC 411
 Db 66 GlyProCysHisPheValAsnGlyAspSerThrProSerLeuAsnGluAsnSerTyrPsn 85
 QY 412 GAGGTCTCCAAATTTACTATCTGTCCCGCATTTGGAGTCGGGCTTTTCATATAGTAT 471
 Db 86 AsnTyrAlaAsnMetIleTyrIleAspGlnProIleGlyValGlyPheSerTyrGly--- 104
 QY 472 ACGGTGTAGTGGTCCATTAACTGTAACCTGTAACCTGGGTGTCGAAATTCGAGCTTTGCAGGA 531
 Db 104 ----- 104
 QY 532 GTTCAGGCGCGGTACCCAAACCATGATGCCACTCTGATCGATACCTACCAATCTTGGCGCA 591
 Db 105 -----ThrAspValThr-----SerThrValThrAlaAla 115

RESULT 3

US-08-309-341-2

; Sequence 2, Application US/080309341

; Patent No. 5594119

QY 592 GAGCCCGCTTGGGAGATCTCTGCAAGGATTCCTTAGTGGACTACCTAGCTTGGACTCTAGG 651
 Db 116 ProTyrValTyrAsnLeuLeuGlnAlaPheTyrAlaGlnArgProGluTyrGluSerArg 135
 QY 652 GTGCAGCTAAGGACTTCACTCTATGACGAGAGCTATGAGGCGCATATGTCCTGCTGCA 711
 Db 136 -----AspPheAlaIlePheThrGluSerTyrGlyGlyHisTyrGlyProGlu 151
 QY 712 TTCTTCAATCATTTTACGAGCAGAAATGAGAGAAATGCCAACGGTACTGTGTAATGGTGT 771
 Db 152 PheAlaSerTyrIleGluGlnGlnAsnAlaAlaIleLysAlaGlySerValThrGlyGln 171
 QY 772 CAGCTTAATTTCAACTCTCTGGGAATTTAATTAACGGCATCATCGACGAGGCGATCCAGGCC 831
 Db 172 AsnValAsnIleValAlaLeuGlyValAsnAsnGlyTyrPheAspSerThrIleGlnGlu 191
 QY 832 CCTTACTACCTGAATTCGCTGTGAACATACCTACGGTATCAAGGCTGTCAACGAGACC 891
 Db 192 LysAlaTyrIleAspPheSerTyrAsnAsnSerTyrGlnGlnIleIleAspSerSerThr 211
 QY 892 GTCTACAATCTACATGAAGTTTGCAACCAATGCCAAATGGTTCGCCAGGATTTGATTCC 951
 Db 212 ArgAspSerLeuLeuAspAlaTyrAsn-----AsnGlnCysLeuProAlaLeuGln 228
 QY 952 ACCTGCAACAGACAAACCCGACCGCATTAGCTAGCTAGCCCTCTCGCGCGAGGCCACC 1011
 Db 229 GlnCysSerGlnSerGlySerThr-----SerAspCysThrAsnAlaAsp 243
 QY 1012 AACATGTGAGGAGCAATGTTGAGGGGCGCATACTACGCCCTTCTGCTGGTGGTGTAT 1071
 Db 244 SerValCysTyrGlnAsnIleGluGlyProIleSerSerSerGlyAspPheAspValTyr 263
 QY 1072 GATATTGGCATCTCATATGATGACCCGACTCCGCGCAAGTTATTACAACTTTCTGGCA 1131
 Db 264 AspIleArgGluProSerAsnAspProTyrProTyrThrTyrSerThrTyrLeuSer 283
 QY 1132 AAGGACTCTGTATGAGACGCTATCGCGGTCACATCACTACACCCAGCTCCCAATATGAC 1191
 Db 284 AspProThrValValIleAlaIleGlyAlaArgThrAsnTyrGlnGluCysProAsnGly 303
 QY 1192 GTCTACTAGCTTCCAGCAACAGCGGACTTTGTGCGCCCACTTCATCGAAGACCTC 1251
 Db 304 ProTyrAsnLysPheAlaSerThrGlyAspAsnPro---ArgSerPheLeuSerThrLeu 322
 QY 1252 GAGGAGATCTTGTCTCTCCCGTGGGTGTCTCCCTCATCTATGCGGAGCCGCTATCATC 1311
 Db 323 SerSerValValGlnSerGlyIleAsnValLeuValTyrAlaGlyAspAlaAspTyrIle 342
 QY 1312 TGCAACTGGTTCGGCGGTACAGCCGCTTTCCCTCGCTGCGAACTACTCCCAAGCCGCCAG 1371
 Db 343 CysAsnTyrLeuGlyAsnTyrGluValAlaAsnAlaValAspPheProGlyAsnAlaGln 362
 QY 1372 TTCGAGCCGAGGTACACGCCCTTGAAAGTCAACGGCGTCGAGTATGGGAACTCGC 1431
 Db 363 PheSerAlaLeuAspLeuAlaProTyrThrValAsnGlyValGluLysGlyGlnPheLys 382
 QY 1432 GAGTATGTAATTTCTCTCTCTACCTCGCTCTATGGCAGCGCCATGAGTCCATCTAC 1491
 Db 383 ThrValAspAsnPheSerPheLeuLysValTyrGlyAlaGlyHisGluValProTyrTyr 402
 QY 1492 CAGCCCATCGCTCTCTCTGCAATTTTAAACCGGACTATCTTCGGTGGGATATCGCAGAG 1551
 Db 403 GlnProAspThrAlaLeuGlnAlaPheLysGlnIleIle----- 415
 QY 1552 GGCCAGAGAGAG 1563
 Db 416 ---GlnLysLys 418

US-09-712-338-1 (1-1668) x US-08-309-341-2 (1-557)

QY	133	GCAACAAATGTCACCATCCGGTACAAGAACCCGGGCAGCAGGCGCTCGCAGACTACC	192
		: :	
Db	132	AlatyrAspLeuArgValLysThrAspProGlySerLeuGlyIle-----Asp	148
		: :	
QY	193	CCGGGTGTCAAATCCTACTCTGGATATGTCACACCTCTCCC---GAGTCCCATACCTTC	249
		: :	
Db	149	ProGlyValLysGlnTyThrGlyTyLeuAspAsnGluAsnAspLysHisLeuPhe	168
		: :	
QY	250	TTCGTGTTCTCGAAGCCAGACATAACCCAGAACTGCATCATCATTTGTGTTGAAT	309
		: :	
Db	169	TyTrpPheGluSerArgAsnAspProGluAsnAspProValLeuTriPheuAsn	188
		: :	
QY	310	GCTGGCCTCGAAGCGATTCTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTGCATGTC	369
		: :	
Db	189	GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile	208
		: :	
QY	370	AATTTCGACTTTTGATGACTACATCAACCTACTCGTGGACAGAGTCTCCAATTACTA	429
		: :	
Db	209	AsnLysLysIleGlnProValTyrrAsnaAspTyrrAlatrPAsnSerAsnaAlaSerValile	228
		: :	
QY	430	TTCCTGTCCCGACCATTTGGGAGTCGGCTTTTCATATAGTAGTACCGGTTGATGGGTCCATT	489
		: :	
Db	229	PheLeuAspGlnProValAsnValGlyTyrrSerTyrrSerAsnSerAla-----	244
		: :	

Db 552 LeuGlyGlyGluTrp 556

RESULT 4

US-08-608-267-2

Sequence 2, Application US/08608267

Patent No. 5688663

GENERAL INFORMATION:

APPLICANT: Xaver, Debbie Sue

APPLICANT: Thompson, Sheryl Ann

TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5688663o No. 5688663disk of No. 5688663th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/608,267

FILING DATE: 28-FEB-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/309,341

FILING DATE: 20-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lowney, Karen A.

REGISTRATION NUMBER: 31,274

REFERENCE/DOCKET NUMBER: 4247.000-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Aspergillus Niger

US-08-608-267-2

Alignment Scores:

Pred. No.:	2.13e-51	Length:	557
Score:	630.00	Matches:	154
Percent Similarity:	48.45%	Conservative:	81
Best Local Similarity:	31.75%	Mismatches:	174
Query Match:	20.81%	Indels:	76
DB:	1	Gaps:	17

US-09-712-338-1 (1-1668) x US-08-608-267-2 (1-557)

QY 133 GCAACAAATGTCCATCCGGTACAGGAACCCGGGCGAGGGCGTCTGGAGACTACC 192

Db 132 AlaTyrAspLeuArgVallylsysThrAspProGlySerLeuGlyle 148

QY 193 CCGGTGTCAATCTCTGATATGTGACACTCTCC--GAGTCCCATACCTTC 249

Db 149 ProGlyVallylsGlnFyThrGlyTyrLeuAspAspAsnGluAsnAspHisLeuPhe 168

QY 250 TTCTGGTCTTCAGACCCAGACATACCCAGAACTGACCTATCACATTGTGGTTGAAT 309

Db 169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188

QY 310 GTGGCCCTCGAAGCGATTCTTTGATCGTCTCTTCGAGAGTTGGGCCCTTGCCAGTGC 369

Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile 208

QY 370 AATTCGACTTTTGATGACTACATCAACCTCGTGAACGAGGTCTCCAAATTTACTA 429

Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValle 228

QY 430 TTCCTGTCCAGCCATCGGAGTGGCTTTTCATATAGTAGTACAGGTGTGATGGTCCATT 489

Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla----- 244

QY 490 AACCTGTAACTGGGGTCTGCGAAAATTCGAGCTTTCGAGAGTTCAGGCGCGGTACCCA 549

Db 244 ----- 244

QY 550 ACCATTGATGCCACTCTGATCGATACCAATCTTGCCGCGAGAGCGCGTTGGGAGATC 609

Db 245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258

QY 610 CTGCAAGGATTCTTAGTGAGCTACCTAGCTTGGAGCTCTAGGGTGGAGTCTAAGGACTTC 669

Db 259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274

QY 670 AGTCTATGACGGAGAGCTATGAGAGGCATATGCTGCTGCTGCTCAATTCATTTTAC 729

Db 275 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 294

QY 730 GAGCAGATGAGAGAATTCGCAACGGTAGTGTAAATGGTGTCTGAGCTTAATTTCAACTCT 789

Db 295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305

QY 790 CTGGGAATTATTAAACGCATCATCGAGAGCGCATCCAGCCCTTACTACCTGGAATTC 849

Db 306 ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet 325

QY 850 GCTGTGAACAATACCTAGGTATCAAGGCTGTCAAGAGAGCCCTTACAACTACATGAAG 909

Db 326 AlaCysGlyAspGlyGlyTyrProAlaValLeuAspGluSerSerCysGlnSerMetAsp 345

QY 910 TTGTCACCAACCAATGCTTGCAGAGATTGATTTCCACCTGCAACAGACAGAAC 969

Db 346 -----AsnAlaLeuProArg--CysGlnSerMetIleGluSerCysTyrSerSerGlu 362

QY 970 CGCACCGCATTAGCTAGCTAGCCCTCTGCGGAGAGCCACCAACATGTGAGGAGCAAT 1029

Db 363 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 377

QY 1030 GTTGAGGGCCCATACTACGCCCTTGTGCTGCTGCTGCTGATGATATTCGGCATCCATAT 1089

Db 378 LeuLeuAlaProTyrGlnArg--ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396

QY 1090 GATGACCCG-----ACTCGCCCAAGTTATTACACAAATTTCTGGCAAG 1134

Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416

QY 1135 GACTCTGTCATGACGCTATCGGCTCAACATCAAC---TACACCCAGTCCCAATATGAC 1191

Db 417 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436

QY 1192 GTCTACTACGCTTTCAGCAACAGCGACTTGTCTGGCCC-----AATTCTATCGAA 1245

Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 456

QY 1246 GACTCGAGGAGATCTTGTCTCTCCCGTGGCTGTCTCCCTCATCTAT---GGCGACCC 1302

Db 457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471

QY 1303 GATTACATCTCAACTGTTTCGGCGGTGAGCGCTTCCCTCGCTGCGAATCTACTCCCAA 1362

Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTyrProGly 491

QY 1363 GCCGCCCATGTCGAGAGCGGTACACGCCCTCTGAAGTCT-----AAC 1407

Db 492 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 511

1408 GGCTGAGTATGGGAACTCGGAGTATGTAATTTCTCTTCACTCGCTATGAG 1467
149 ProGlyValIysGlnIleThrGlyLeuAspAsnGluAspHisLeuPhe 168
250 TTCTGGTCTTGAAGCAGACATACCCAGAACTGACCTATCATCTGTTGTTGAAT 309
169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188
310 GGTGGCTGGAAGCGATTCTTTCATCGCTCTCTCGAAGAGTTGGGCGCTTGCATCTC 369
189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
370 AATTGCGATTTTGTGATGACTACATCAACCTCTGCTGGAACAGAGTCTCCAAATTA 429
209 AsnLysLysIleGlnProValIleThrValIleThrValIleThrValIleThrValIle 228
430 TTCTGTCTCCAGCCATGGGAGTGGCTTTTCATATAGTATGATACGGTTGATGGTCCAT 489
229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla 244
490 AACCTGTAACTGGGCTGTCGAAAATTCGAGCTTTGCGAGGATTCAGGGCGGTACCCA 549
244 ----- 244
550 ACCATTGATGCCACTCTGATGATACCTACATCTTGGCGAGAGCGCGTGGAGATC 609
245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
610 CTGCAAGGATTCCTTAGTGAGTACTAGCTTGGACTCTAGGTGAGTCTAAGGACTTC 669
259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274
670 AGTCTATGACCGAGAGCTATGAGGAGCTATGTCCTGCTATCTTCAATCATTTTAC 729
275 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 294
730 GAGCAGATGAGAGAAATTCGCAACGCTAGTGTATGTTGTTGCTTCAATTTCACTCT 789
295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305
790 CTGGGAATTTAATTAACGGCATTCAGAGAGGAGATCCAGGCCCTTACTACCTGATTC 849
306 ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGlyTyrArgProMet 325
850 GCTGTGAAACAATACCTACGGTATCAAGGCTGTCAACGAGACCGCTTCAACTACATCAAG 909
326 AlaCysGlyAspGlyGlyTyrProAlaValLeuAspLysSerSerCysGlnSerMetAsp 345
910 TTTCGCAACCAATGCAATGTTGCGAGGATTTGATTTCCACCTGCAACAGACAAAC 969
346 -----AsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
970 CGCACCGCATTAGCTAGTACGCCCTCTGCGCGAGACCCACCAACATGTCAGGGACAAT 1029
363 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 377
1030 GTTCAGGGCCATACCTAGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
1090 GATGACCCG-----ACTCGCCAAAGTTATTACAAATTTCTCGCAAG 1134
397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
1135 GACTCTGTCATGAGCGTATCGGCTCAACATCAAC---TACACCCAGTCCAAATATGAC 1191
417 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
1192 GTCTACTAGCTTTCCAGCAACAGCGGAGCTTTGTCTGGGCC-----AACTTCATCGAA 1245
437 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 456
1246 GACCTCGAGGAGATCTTGTCTCTCCCGCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1302

QY 1408 GGCTGAGTATGGGAACTCGGAGTATGTAATTTCTCTTCACTCGCTATGAG 1467
Db 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
QY 1468 GCAGGCCATGAAGTCCCATACCTACAGCCCATCGCTCCCTGCAATTTTAACCGGACT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556
RESULT 5
US-08-608-452-2
; Sequence 2, Application US/08608452
; Patent No. 5693510
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5693510o No. 5693510disk of No. 5693510th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,452
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
US-08-608-452-2
Alignment Scores:
Pred. No.: 2,13e-51 Length: 557
Score: 630.00 Matches: 154
Percent Similarity: 48.45% Conservative: 81
Best Local Similarity: 31.75% Mismatches: 174
Query Match: 20.81% Indels: 76
DB: 1 Gaps: 17
US-09-712-338-1 (1-1668) x US-08-608-452-2 (1-557)
QY 133 GCAACAATGTCACCATCGGTACAAAGGAACCCGGGCGAGAGGGCGTCTGCGAGACTACC 192
Db 132 AlaTyrAspLeuArgValIysLysThrAspProGlySerLeuGlyIle-----Asp 148

Db	457	GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla	471
Qy	1303	GATTACATCTGCAACTGTTCCGGCGGTACAGCCGCTTCCCTCGCTCGCACTACTCCCAA	1362
Db	472	AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly	491
Qy	1363	CGCGCCCATGTTCCGAAGCGCAGGTACACGCCCTCGAAAGTC-----AAC	1407
Db	492	GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr	511
Qy	1408	GGCTCGAGTATGGGAAACTCGCGAGTATGGTAATTTCTCCTTCATCTCGCTCATGAG	1467
Db	512	GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly	531
Qy	1468	GCAGGCCATGAAGTCCCATCTACTACCGCCCATCGCTCCCTGCAATGTGTTACCGGACT	1527
Db	532	GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp	551
Qy	1528	ATCTTCGGT---TGG	1539
Db	552	LeuGlyGlyGluTrp	556

RESULT 6

```

US-08-608-224-2
; Sequence 2, Application US/08608224
; Patent No. 5705376
;
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.
; STREET: 405 Lexington Avenue, 3rd Floor
; CITY: New York, N.Y. 10017-2498
; COUNTRY: U.S.A.

```

```

/ / COMPUTER READABLE FORM:
/ / MEDIUM TYPE: Floppy disk
/ / COMPUTER: IBM PC compatible
/ / OPERATING SYSTEM: PC-DOS/MS-DOS
/ / SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
/ / CURRENT APPLICATION DATA:
/ / APPLICATION NUMBER: US/08/608,224
/ / FILING DATE: 28-FEB-1996
/ /

```

Alignment Scores:		
red. No.:	2.13e-51	Length: 557
Score:	630.00	Matches: 154
Percent Similarity:	48.45%	Conservative: 81

Best Local Similarity:	31.75%	Mismatches:	174
Query Match:	20.81%	Indels:	76
DB:	1	Gaps:	17
US-09-712-338-1 (1-1668) x US-08-608-224-2 (1-557)			
QY	133	GCAAACAATGTACCATTCCGGTACAAAGGAACCCGGGCGCAGAGGGCGTCTCGGAGACTACC	192
Db	132	AlaTyrAspLeuArgValIysLysThrAspProGlySerLeuGlyIle-----Asp	148
QY	193	CCGGGTCTCAATCCTACTCTGATATATGACACCTCTCCC---GAGTCCCATCTACTTC	249
Db	149	ProGlyValIysGlnTyrThrGlyTyrLeuAspAspAsnGluAsnAspLysHisLeuPhe	168
QY	250	TTCTGTGTTCTTCGAAGCCAGACATACCCAGAACTGCACCTATCACATTGTGGTTGAAT	309
Db	169	TyrTriPhePheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn	188
QY	310	GGTGGCCCTGGAGCGAATCTTTTGATCGCTCTCTCGAAGAGCTTGGGCCCTTGCCATGTC	369
Db	189	GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle	208
QY	370	AATTCGACTTTTGATGACTACATCAACCTCTACTCGTGAAGAGTCTCCAATTACTA	429
Db	209	AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle	228
QY	430	TTCTGTCCAGCCATTCGGAGTCGGCTTTTCATATAGTGATACGGTTGATGGTCCATT	489
Db	229	PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----	244
QY	490	AACCTGTTAACGTGGGGTCTGCGAAATTCGAGACTTTGAGGAGTTTCAGGCGCGGTACCCA	549
Db	244	-----	244
QY	550	ACCATTGATGCCACTCTGATCGACTACTACCAATCTTCGCCGAGAGGCGGTGGAGATC	609
Db	245	-----ValserAspThrValAlaAlaGlyLysAspValTyrAlaLeu	258
QY	610	CTGCAAGGATTCCTTAGTGGACTACCTAGCTTGGACTCTAGGCTGACGTCCTAAGGACTC	669
Db	259	LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe	274
QY	670	AGTCTATCGAGCGAGACTATGGAGGCATATGGTCTCTCATTTCTTCAATCATTTTATC	729
Db	275	HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu	294
QY	730	GAGCAGAATGAGAGAATTCGCAACGGTAGTGTTAATGGTGTTTCAGCTTATTTCACTCT	789
Db	295	SerHisLysLysArg-----AsnIleAsnLeuGlnSer	305
QY	790	CTGGGANTATTAAACGGCATCATCGACGAGCGCATCCAGGCCCTTACTACCTGATTC	849
Db	306	ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet	325
QY	850	GCTGTGCAACAATACCTACGGTATCAAGGTGTCAACGACCGTCTPACAACTACATGAAG	909
Db	326	AlaCysGlyAspGlyTyrProAlaValLeuAspGluSerSerCysGlnSerMetAsp	345
QY	910	TTTGCCAAACCAATGCCAAATGGTTCGCCAGGATTTGATTTCACCTGCAACAGACAAC	969
Db	346	-----AsnAlaLeuProArg--CysGlnSerMetIleGluSerCysTyrSerSerGlu	362
QY	970	CGCACCGCATTAGCTACTACGCCCTCTCGCCGCAAGCCACCAATGTGAGGAGCAAT	1029
Db	363	SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla	377
QY	1030	GTTGAGGGCCACTACTACGCCCTTCTGGTCTGTGTGTATGATATTCGGCATCCATAT	1089
Db	378	LeuLeuAlaProTyrGlnArg--ThrGlyGlnAsnValTyrAspValArgGlyLysCys	396
QY	1090	GATGACCCCG-----ACTCCGCCAAGTTATTACAAATTTCTGGCAAG	1134
Db	397	GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys	416

1135	GA	CTCTG	CATG	CAGC	GC	TATCG	CGTCA	ACATCAAC	----	TAC	CC	CAG	TCC	ATA	TGAC	1191					
417	Pro	Leu	Val	Ile	Glu	Ala	Gly	Ala	Glu	Val	Asn	Gly	Tyr	Asp	Ser	Cys	Asn	Phe	Asp	436	
1192	GT	CTACT	AC	CGT	TTCC	AGCA	AA	CAG	CG	ACTT	GT	CTG	GGCC	-----	AAC	TTCA	TC	GAA	1245		
	::																				
437	Ile	Asn	Arg	Asn	Phe	Leu	Phe	His	Gly	Asp	Pro	Met	Lys	Pro	Tyr	His	Arg	Leu	Val	Pro	456
1246	GAC	CTC	GAG	GAG	ATC	CTT	TGCT	CTC	CCG	TGCG	TCT	CCCT	CTCAT	CTAT	----	GGC	GAC	GC	1302		
457	Gly	Leu	Leu	Glu	-----	Gln	Ile	Pro	Val	-----	Leu	Ile	Tyr	Ala	Gly	Asp	Ala	471			
1303	GAT	TAC	AT	TCG	AACT	GGT	CGG	CGT	CAG	CGG	TTT	CCCT	CGCT	CGCA	ACTACT	CCCA	1362				
472	Asp	Phe	Ile	Cys	Asn	Trp	Leu	Gly	Asn	Pro	Leu	Glu	Ala	Leu	Glu	Trp	Pro	Gly	491		
1363	GCC	GCC	CAG	TGCC	GAAG	CGC	GAG	GTAC	AC	GC	CCCT	CGAA	GTC	-----	AAC	1407					
492	Gln	Ala	Glu	Tyr	Ala	Ser	Ala	Glu	Leu	Glu	Asp	Leu	Val	Ile	Val	Asp	Asn	Glu	His	Thr	511
1408	GCG	TCG	AGT	ATG	GGG	AA	CTCG	GAG	TAT	CGT	ATG	TAT	TCT	CCCT	CAC	T	CGG	CTCAT	TGAG	1467	
512	Gly	Lys	Ile	Gly	Gln	Val	Lys	Ser	His	Gly	Asn	Phe	Thr	Phe	Met	Arg	Leu	Tyr	Gly	531	
1468	GC	AGC	CCAT	GAAG	TCCC	ATA	CTACT	TACC	AGCC	CACT	CGC	CTCC	CTCG	CAAT	TGTTT	TAA	CCG	GACT	1527		
532	Gly	Gly	His	Met	Val	Pro	Met	Asp	Gln	Pro	Glu	Ser	Ser	Leu	Glu	Phe	Phe	Asn	Arg	Trp	551
1528	AT	CTT	CGT	-----	TGG	1															

RESULT 7

US-08-967-149-2
Sequence 2, Application US/08967149
Patent No. 5939305
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
TITLE OF INVENTION: ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59393050 No. 5939305disk of No. 5939305th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,149
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,452
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:


```

Db 346 -----AsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
QY 970 GCGACCGCAATAGCTAGCTAGCCCTCTGCGCGGAGCAACCAACATCTGCGAGGACAAAT 1029
Db 363 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAla 377
QY 1030 GTTCAGGGGGCATACTAGCCCTTCTGCTGGTGTGTATGATATTCGGCATCCATAT 1089
Db 378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
QY 1090 GATCACCGC-----ACTCCGCCAGTTATTACAAACAATTTCTGGCAAG 1134
Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
QY 1135 GACTCTGTCTGACGCTATCGGGGTCAACATCAAC---TACACCCAGTCCCAATAATGAC 1191
Db 417 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
QY 1192 GTCTACTACGCTTTCACGACAAACAGGACGCTTCTGTGGGCC-----AATCTTCATCGAA 1245
Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 456
QY 1246 GACCTCCGAGGAGATCTTCTCTCCCGTGGTGTCTCCCTCATCTAT---GGCGAGGCC 1302
Db 457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471
QY 1303 GATTACATCTGCACTGTTTCGCGGTCTGCGGCTGAGCGGCTTTCCTCTGCGAACTACTCCAA 1362
Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 491
QY 1363 GCGCGCCAGTTCGAGCGGAGGTCACGCGCTGAAAGTC-----AAC 1407
Db 492 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 511
QY 1408 GCGCTCGAGTATGGGAAACTCGCGAGTATGTAATTTCTCTCTACTCGCTCTATGAG 1467
Db 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
QY 1468 GCAGGCAATGAATCCCATCTACTACGAGCCATCGCTCCCTGCAATTTTAAACCGGACT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnA-gTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556

```

RESULT 8

```

US-08-309-341-4
; Sequence 4, Application US/08309341
; Patent No. 5594119
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLIUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSES: No. 55941190 No. 5594119disk of No. 5594119th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/309,341
; FILING DATE: 16-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.

```

```

; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247,000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-309-341-4

```

Alignment Scores:

```

Pred. No.: 9,94e-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 1 Gaps: 18

```

US-09-712-338-1 (1-1668) x US-08-309-341-4 (1-557)

```

QY 133 GCAACAATCTACCATTCGGTACAGGAACCCGGGCGAGAGGCGTCTCGAGACTACC 192
Db 132 AlatyAspLeuArgValLysLysThrAspProSerSerLeuGlyLys-----Asp 148
QY 193 CCGGGTCTCAATCTCTACTCTGATATGTCGACCTCTCCC---GAGTCCCATCTTC 249
Db 149 ProGlyValLysGlnTyrThrGlyTyrLeuAspAsnGluAsnAspLysHisLeuPhe 168
QY 250 TTCTGTCTTCGAAGCGATTCCTTGTATCGGTCTCTCGAAGAGTTGGCCCTTCCCATGTC 369
Db 169 TyrTyrPhePheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188
QY 310 GTTGGCCCTTGAAGCGATTCCTTGTATCGGTCTCTCGAAGAGTTGGCCCTTCCCATGTC 369
Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
QY 370 AATTCGACTTTTGATGACTACATCAACCTCTCTCTGGAACGAGTCTCCAAATTTACTA 429
Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle 228
QY 430 TTCTGTCTCCAGCCATTCGGAGTTCGGCTTTTCATATAGTATGATACGCTTATGGTCCATT 489
Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----244
QY 490 AACCTGTAACTGGGGTCTGTCGAAAATTCGAGCTTTCGAGGAGTTTCAGGGCGGTACCCA 549
Db 244 -----244
QY 550 ACCATTGATGCCACTCTGATCGATACCTACCAATCTTGGCGAGAGCGCGCTTGGGAGATC 609
Db 245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
QY 610 CTGCAAGGATTCCTTAGTGGACTACCTAGCTTGGACTCTAGGTGCTAGTCTTAAGGACTTC 669
Db 259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274
QY 670 AGTCTATCGAGAGAGCTATGAGGGGCACTATGCTCTGATCTTCAATTCATTTTAC 729
Db 275 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluLeu 294
QY 730 GAGCAGATGAGAAATTCGCAACGGTAGTGTAAATGTGTTCAGCTTAATTTCAACTCT 789
Db 295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305
QY 790 CTGGGAATTATTAACGGCATCATCGACGAGCGGATTCAGGCCCTTACTACCTGAATTC 849
Db 306 ValLeuIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGluTyrArgProMet 325

```

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/608,267
FILING DATE: 28-FEB-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger

US-08-608-267-4

Alignment Scores:
Pred. No.: 9,948-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 18 Gaps: 1

US-09-712-338-1 (1-1668) x US-08-608-267-4 (1-557)

QY 133 GCAACAAATGTCCACATCCGATACAGGAACCCGGGCGAGAGGCGCTCTCGAGATACC 192
Db 132 AlaTyrAspLeuArgValLysThrAspProSerSerLeuGlyLeu-----Asp 148
QY 193 CGGGTGTCAAAATCCTACTCTGGATATGTCGACACCTCTCCC---GAGTCCCATACCTTC 249
Db 149 ProGlyValLysGlnTyrThrGlyTyrLeuAspAsnGluAsnAspLysHisLeuPhe 168
QY 250 TTCTGGTCTTCCGAACCCAGACATACCCAGAACTGCACCTATCATCTGCTTGAAT 309
Db 169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValLeuLeuTrpLeuAsn 188
QY 310 GGTGGCCCTGGAAGCGATCTTTGATCGGTCTCTTCGAGAGTGGGCCCTTCCCATGTC 369
Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
QY 370 AATTCCGATTTTGATGACTACATCAACCTCCTCATCTGTTGGAACAGGTCTCCAATTACTA 429
Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValle 228
QY 430 TTCTGTCCCGACCATTTGGGAGTCGGCTTTTCATATAGTATGATGATGATGATGATGATG 489
Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----244
QY 490 AACCCCTGTAACGTGGGTCTGTCGAAATTCGAGAGTTCGAGGCGGTGACCCCA 549
Db 244 -----244
QY 550 ACCATTGATGCCACTCTGATCGATACCTACCAATCTTCCGCGAGAGCGCGTGGGAGATC 609
Db 245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
QY 610 CTGCAAGGATTCCTTAGTGAGTACTACCTAGCTTGGACTCTAGGGTGCAGTCTTAAGACTTC 669
Db 259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274
QY 670 AGTCTATGACGAGAGAGCTATGAGGCGCACTATGGTCTCTGATCTTCAATCAATTTTAC 729

QY 850 GCTGTGAACAATACCTACGGCTCAAGGCTGTCAACGAGACCGCTCTCAACTACATCAAG 909
Db 326 AlaCysGlyAspGly---GlyTyrProAlaValLeuAspGlu---GlySerCysGlnAla 343
QY 910 TTTGCCCAACAAATGCCAAATGGTTGCCAGGATTTGATTTCCACTGCAAAACACACAAC 969
Db 344 MetAspAsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
QY 970 GCGACCGCATTAGCTAGCTACGCTCTGCGCCGAGCCCAACCAATGTCGAGGAGCAAT 1029
Db 363 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAla 377
QY 1030 GTTGAGGGCCATACCTACGCTTTCGCTGCTGTGTGTATGATATTCGGCATCCATAT 1089
Db 378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
QY 1090 GATGACCCG-----ACTCGGCCAAGTTATTACAAATAATTTCTGCAAG 1134
Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
QY 1135 GACTCTGTCATCGAGCTATCGCGCTCAACATCAAC---TACACCCAGTCCCAATATGAC 1191
Db 417 ThrGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
QY 1192 GTCTACTACGCTTCCAGCAACAGCGGACTTTGTCTGCGCC-----AACTTCATCGAA 1245
Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTyrMetLysProTyrHisArgLeuValPro 456
QY 1246 GACCTCGAGAGATCCTGCTCTCCCGTGGGTCTCCCTCATCATAT---GGCGACGCC 1302
Db 457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471
QY 1303 GATTACATCTGCACTGCTGCGCTGCGGCTGCGGCTTCTGCTGCGCACTACTCCCAA 1362
Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 491
QY 1363 GCGGCCAGTTCGCAAGCGGAGGTATACGCGCCCTCGAAAGTC-----AAC 1407
Db 492 GlnAlaGluTyrAlaSerAlaLysLeuGluAspLeuValValGluAsnGluHisLys 511
QY 1408 GCGCTCGAGTATGGGAACTCGGAGTATGGTATTTCTCTCTACTCGCGTCTATGAG 1467
Db 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
QY 1468 GCAGGCCATGAGTCCCATACTACCGCCCATCGCTCTCCGCAATTTGTTAACCGGACT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556

RESULT 9
US-08-608-267-4
Sequence 4, Application US/08608267
Patent No. 5688663
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5688663o No. 5688663disk of No. 5688663th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

Db 275 HsIleAlaGlySerTyrAlaGlyHisTyrIleProValPheAlaSerGluLeuLeu 294
QY 730 GAGCAGAATGAGAGAAATCCCAACGGTAGTGTAAATGGTTGTTCACTTAATTTCAACTCT 789
Db 295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305
QY 790 CTGGGAATTAATACGGCATCATCAGACGGCGATCCAGGCCCTTACTACCCCTGAATTC 849
Db 306 ValLeuIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGluTyrArgProMet 325
QY 850 GCTGTGAACAATACCTACGGTATCAAGCTGTCAACGACCGCTCAACAACACTACATGAAG 909
Db 326 AlaCysGlyAspGly---GlyTyrProAlaValLeuAspGlu---GlySerCysGlnAla 343
QY 910 TTGTGCAACCAATCCCAATAGTTTGCAGGATTTGATTTCCACCTGCAAAACAGACAAC 969
Db 344 MetAspAsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
QY 970 CGCACCAGCTAGCTAGCTAGCCCTCTGCCCGCAGCCACCAACATGTGACGAGGACAAT 1029
Db 363 SerAla-----TyrValCysValProAlaSerIleTyrCysAsnAsnAla 377
QY 1030 GTTCAGGGCCCATACCTAGCCCTTGTCTGCTGCTGTGTGTATGATATTCGGCATCCATAT 1089
Db 378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
QY 1090 GATCACCAG-----ACTCCGCAAGTTATTACAAACAAATTTCTGGCAAG 1134
Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
QY 1135 GACTCTCTCATGAGCGTATCGGGCTGCACATCAAC---TACACCCAGTCCCAATATGAC 1191
Db 417 ThrGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
QY 1192 GTCTACTACGCTTTCCAGCAACACAGCGACTTTCTCTGGGCC-----AACTTCATCGAA 1245
Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTyrMetLysProTyrHisArgLeuValPro 456
QY 1246 GACTCTGAGGAGATCTTGTCTCCCGCTGCTGCTCTCTCATCTAT---GGCGAGCC 1302
Db 457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471
QY 1303 GATTACATCTGCAACTGTTGCGGGTGCAGCGCTTCCCTCGCTGCACTACTCCCAA 1362
Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 491
QY 1363 GCGGCCAGTTCCGAAGCGCAGGGTACACGCCCTCGAAGTC-----AAC 1407
Db 492 GlnAlaGluTyrAlaSerAlaLysLeuGluAspLeuValValGluAsnGluHisLys 511
QY 1408 GCGTCGAGTATGGGAAACTCCGAGTATGTAAATTTCTCTTCACTCGCGTCTATGAG 1467
Db 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
QY 1468 GCAGGCCATGAGTCCCATCTACAGCCCATGCGCTCCCTGCAATGTTTAAACCGGACT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556

```

RESULT 10

```

US-08-608-452-4
; Sequence 4, Application US/08608452
; Patent No. 5693510
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: No. 5693510o No. 5693510disk of No. 5693510th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,452
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-608-452-4

```

Alignment Scores:

```

Pred. No.: 9,94e-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 1 Gaps: 18

```

```

US-09-712-338-1 (1-1668) x US-08-608-452-4 (1-557)

```

```

QY 133 GCAAAACAATGTCAACATCGGTACAGGAACCGGGCAGAGGGCGTCTCGGAGACTACC 192
Db 132 AlaTyrAspLeuArgValLysLysThrAspProSerSerLeuGlyIle-----Asp 148
QY 193 CCGGFTGTCAAAATCCTACTCTCGATATGTCGACACCTCTCCC---GAGTCCCATACCTTC 249
Db 149 ProGlyValLysGlnTyrThrGlyTyrLeuAspAsnGluAsnAspLysHisLeuPhe 168
QY 250 TTCTGGTCTTCGAAGCCAGACATAACCCAGAACTGCACCTATCACATTGTGGTGAAT 309
Db 169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValLeuTrpLeuAsn 188
QY 310 GGTGGCCCTGGAAGCGATTCTTGTATCGTCTCTTCGAGAGAGTTGGGCCCTTGCCTATC 369
Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
QY 370 AATTGACGCTTTTGATGACTACATCAACCTCACTCGTGAACAGAGGTCTCCCAATTACTA 429
Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle 228
QY 430 TTCCTGTCCCGCAGCATGGGAGTCGGCTTTTCATATAGTATACGTTGATGGGTCCATT 489
Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----244
QY 490 AACCTGTACTGGGTCGTCGAAAAATTCGAGGCTTTCGAGGAGTTTCAGGCGCGTACCCA 549
Db 244 -----244

```

```

RESULT 11
US-08-608-224-4
; Sequence 4, Application US/08608224
; Patent No. 5705376
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,224
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-608-224-4

Alignment Scores:
Pred. No.: 9,94e-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 1 Gaps: 18

US-09-712-338-1 (1-1668) x US-08-608-224-4 (1-557)
Qy 133 GCRAACAATGTCACCATCCGGTACAGAAACCCGGGACAGAGGGGCTCTGCGAGACTACC 192
Db 132 AlaTyrAspLeuArgValLysLysThrAspProSerSerLeuGlyIle-----Asp 148
Qy 193 CCGGGTGTCAATCCTACTCTGGATATGTCGACACCTCTCCG---GAGTCCCATCTTC 249
Db 149 ProGlyValLysGlnTyrThrGlyTyrLeuAspAsnGluAsnAspLysHisLeuPhe 168
Qy 250 TTCGTGTTCTTCGAAGCCAGACATACCCAGAACTGCACCTATCACATTCTGGTTGAAT 309
Db 169 TyrTrpPhePheGluSerArgAsnAspProGluAsnAspProValValLeuTriPheAsn 188
Qy 310 GTGTGCCCTCGAAGCGGATTCTTTGATCGGTCTCTTCGAAGAGATTGGCCCTTCCCATGTC 369
Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
Qy 370 AATTCGACTTTTGTGACTATCATCAACCCCTACTCGTGGACGAGTCTCCCAATTTACTA 429
```

Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle 228
QY 430 TTCTGTCCAGCCATTGGAGTCGGCTTTTCATATAGTATACGGTTGATCGGTCCATT 489
Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla 244
QY 490 AACCTGTAACTGGGTGCTGAAAAATTCGAGCTTTGCAGGAGTTTCAGGGCGGTACCCA 549
Db 244 244
QY 550 ACCATTGATGCCACTCTGATGATACCTACCAATCTTGCCGAGAGCGCGTTGGAGATC 609
Db 245 ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
QY 610 CTCAAGGATTCCTTAGTGGACTACCTAGCTTGAGCTCTAGGGTCGAGCTTAAGGACTTC 669
Db 259 LeuThrLeuPhePheLysGlnPhePro 274
QY 670 AGTCTATGGAGGAGACTATGGAGGACTATGCTCTGCAATCTTCAATCACTTTTAC 729
Db 275 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 294
QY 730 GAGCAGAAATCAGAAATTCGCAACGGTAGTGTAACTGCTTCACTTAATTTCAACTCT 789
Db 295 SerHisLysLysArg 305
QY 790 CTGGGAATTAATTAACGGCATCTGACGAGCGGATCCAGGCGCTTACTACCTCCGTAATTC 849
Db 306 ValIleIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGluTyrTyrArgProMet 325
QY 850 GCTGTGAACAACTACCTACGCTATCAAGCTGTCAACGAGACCGCTCTCAACTACATCAAG 909
Db 326 AlaCysGlyAspGly 343
QY 910 TTTGCCAACCAATGCAATGCTGCCAGATTTGATTCACCTGCAACAGACAAAC 969
Db 344 MetAspAsnAlaLeuProArg 362
QY 970 CGCACCAGCATTAGCTACGCGCTCTGCGCGGACCAACCAACATGTCGAGGACAT 1029
Db 363 SerAla 377
QY 1030 GTTGAGGGGCCATACCTAGCGCTTTGCTGTGTGTATGATATTCGGCATCCATAT 1089
Db 378 LeuLeuAlaProTyrGlnArg 396
QY 1090 GATGACCGC 1134
Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
QY 1135 GACTCTGTGATGACGCTATCGCGCTCAACATCAAC--TACACCCAGTCCCAATATGAC 1191
Db 417 ThrGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
QY 1192 GTCTACTACGCTTCCAGCAACAGCGGACTTTGTCTGGCC--AACTTCATCGAA 1245
Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTyrMetLysProTyrHisArgLeuValPro 456
QY 1246 GACCTCGAGGAGATCTTCTCTCCCGCTGCTGCTCCCTCATCTAT--GGCAGCGCC 1302
Db 457 GlyLeuLeuGlu 471
QY 1303 GATTACATCTGCAACTGTTTCGGCGGTAGGCGGTTTCCCTCGCTCGGACTACTCCCAA 1362
Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 491
QY 1363 GCCGCCAGTTCGAGCGGCTACACGCCCTCGAAAGTC--AAC 1407
Db 492 GlnAlaGluTyrAlaSerAlaLysLeuGluAspLeuValValGluAsnGluHisLys 511
QY 1408 GCGCTCGAGTAGTGGGAACCTCGGAGTAGTGTAAATTTCTCTTCACTCGGCTCATGAG 1467

Db 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
QY 1468 GCAGGCCATGAAGTCCCATATACAGCCCATCGCTCCCTCGCAATGTTTAAACCGACT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556

RESULT 12

US-08-967-149-4
; Sequence 4, Application US/08967149
; Patent No. 5939305
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
; TITLE OF INVENTION: ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5939305 No. 5939305disk of No. 5939305th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,149
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,452
; FILING DATE: 28-FEB-1996
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger

US-08-967-149-4
Alignment Scores:
Pred. No.: 9,94e-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 2 Gaps: 18

US-09-712-338-1 (1-1668) x US-08-967-149-4 (1-557)
QY 133 GCAACCAATGTCCATCCGTACAGGACCCGGGCGAGGCGCTCGAGACTACC 192
Db 132 AlaTyrAspLeuArgValLysThrAspProSerSerLeuGlyIle-----Asp 148

193 CCGGTTGTCAAATCCTACTCTGGATATGTCGACACCTCTCC---GAGTCCCATACCTTC 249
149 ProglyVallysGlnTyrThrGlyTyrLeuAspAsnGluAsnAspLysHisLeuPhe 168
250 TTCGGTTCTTCGAGCCAGACATACCCAGAACTGACCTATACATCTGCTTGAAT 309
169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValValLeuTyrLeuAsn 188
310 GGTGGCCCTGAGAGGATCTCTTCTGATCGTCTCTGAGAGTGTGGGCCCTTGCATGTC 369
189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile 208
370 AATTCCGACTTTTGATGACTACATCAACCTCTACTCGTGGAGAGGTCTCCAAATTACTA 429
209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTyrAsnSerAsnAlaSerValile 228
430 TTCCTGTCCCGACCATTTGGAGTCGCTTTCATATAGTATGATACGGTTGATGGTCCAT 489
229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnAla----- 244
490 AACCTGTAACTGGGGTCTGTCGAAAATTTCGAGCTTTTCGAGGAGTTTCAGGGCCGCTACCCA 549
244 ----- 244
550 ACCATTGATGCCACTCTGATCGATACACTACCAATTTTGGCGAGAGCCGCTTGGAGATC 609
245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
610 CTGCAAGGATTCCTTGTAGTGACTACCTAGCTTGGACTCTAGGGTGGAGCTCAAGACTTC 669
259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274
670 AGTCTATGAGCGAGAGCTATGAGGCGACTATGCTCTGCAATCTTCAATCATCTTTTAC 729
275 HisIleAlaGlyGlnSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 294
730 GAGCAAGATGAGAGAATTGCCAAGCGTGTGTTTAATGGTTCAGCTTAAATTTCAACTCT 789
295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305
790 CTGGGAATTAATTAAGCATCATCGACGAGCGATCCAGGCCCTTACTACCTCGAATTC 849
306 ValLeuIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGlyTyrArgProMet 325
850 GCTGTGAAACATACCTACGCTATCAAGCTGTCAACGAGACCGCTCTACAACTACATGAAG 909
326 AlaCysGlyAspGly---GlyTyrProAlaValLeuAspGlu---GlySerCysGlnAla 343
910 TTTGCCAACCAATGCGAAATGTTGCCAGATTGATTTCCACCTGCAACACAGACAAAC 969
344 MetAspAsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
970 GGCACCGCATAGCTGACTAGCCCTCTGCGCGAAGCCCAACACATGTCAGGGACAAT 1029
363 SerAla-----TyrValCysValProAlaSerIleTyrCysAsnAla 377
1030 GTTCAGGGGCGCATACAGCCCTTGTGCTGCTGCTGATGATATTCGGCATCCATAT 1089
378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
1090 GATGACCGG-----ACTCCGCAAGTTATTACACAAATTTCTGGCAAG 1134
397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
1135 GACTCTGTGACGCTATCGGCTCAACATCAAC---TACACCCAGTCCAAATGATGAC 1191
417 ThrGluValIleGluAlaValIleGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
1192 GTCTACTACGCTTTCAGCAAAACAGGCGACTTTGTCTGGCC-----AAGTTATCGAA 1245
437 IleAsnArgAsnPheLeuPheHisGlyAspTyrMetLysProTyrHisArgLeuValPro 456
1246 GACCTCGAGGAGATCCTTCTCTCCCGCTGCTCTCCCTCATCTAT---GGCGACGCC 1302

457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471
1303 GATTACATCTGCAACTGGTTCGGCGGTGACGCCCTTTCCCTCGCTCGGAACCTACTCCAA 1362
472 AspPheIleCysAsnTyrLeuGlyAsnLysAlaTyrThrGluAlaLeuGluTyrProGly 491
1363 GCCGCCAGTTCGGAAGCGGAGGTACACGCCCTCGAAAGTC-----AAC 1407
492 GlnAlaGluTyrAlaSerAlaLysLeuGluAspLeuValValGluAsnGluHisLys 511
1408 GCGCTCGAGTATGGGAAACCTCGCAGTATGTAATTTCTCTCTCACTCGCGTCTATGAG 1467
512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
1468 GCAGCCATGAGTCCCATCTACACGCCCATCGCTCCCTGCAATTTCTTAACCGGACT 1527
532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 551
1528 ATCTTCGGT---TGG 1539
552 LeuGlyGlyGluTyr 556
RESULT 13
US-09-640-305-4
Sequence 4, Application US/09640305
Patent No. RE37447
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Yeh, Patrice
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/640,305
FILING DATE: 16-Aug-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/360,673
FILING DATE: 06-FEB-1995
APPLICATION NUMBER: WO PCT/FR93/00623
FILING DATE: 23-JUN-1993
APPLICATION NUMBER: FR 92/07785
FILING DATE: 25-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92040-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-640-305-4
Alignment Scores:

```

Pred. No.: 1,48e-45 Length: 491
Score: 568.50 Matches: 151
Percent Similarity: 44.65% Conservative: 66
Best Local Similarity: 31.07% Mismatches: 184
Query Match: 18.78% Indels: 85
DB: 1 Gaps: 14

US-09-712-338-1 (1-1668) x US-09-640-305-4 (1-491)

QY 133 GCAACAATGTCACATCCCGGTACAAAGAACCCGGGCGAGGGCGTCTGGAGACTACC 192
Db 62 AlatySerLeuArgIleLysProLeuAspProLysSerLeuGlyVal----- 77
QY 193 CCGGCTGTCAAATCCTACTCTGGATATGTCGACACCTCTCCCGAGTCCCATACCTTCTTC 252
Db 78 AspThrValLysGlnTrpSerGlyTyrLeuAspTyrGlnAspSerLysHisPhePheTyr 97
QY 253 TGGTCTTTCGAAGCCACATACCCAGAACTCCACCTATCACATTGCTGTTCAATGGT 312
Db 98 TrpPhePheGluSerArgAsnAspProGluAsnAspProValIleLeuTrpLeuAsnGly 117
QY 313 GGCCTGGAAGCGATTCTTTGATCGGTCTCTTCGAAGATTGGGCCCTTCCCATGTCAAT 372
Db 118 GlyProGlyCysSerSerPheValGlyLeuPhePheGluLeuGlyProSerSerIleGly 137
QY 373 TCAGATTTCGATGACTACATCAACCTCTACTCGTGGAAACGAGGTCTCCAATTACTATTC 432
Db 138 AlaAspLeuLysProIleTyrAsnProTyrSerTrpAsnSerAsnAlaSerValIlePhe 157
QY 433 CTCTCCAGCATTTGGGAGTCGCTTTTCATATAGTATGATACGGTTGATGGTCCATTAAC 492
Db 158 LeuAspGlnProValGlyValGlyPheSerTyrGlyAsp----- 170
QY 493 CTGTAACTGGGTCGTCGAAAATTCGAGCTTTGCAGGAGTTCAGGGCCGGTACCCCAACC 552
Db 170 ----- 170
QY 553 ATGTATGCCACTCTGATCGATACCAATCTTTGCGGAGAGCCGCTTGGGAGATCCTG 612
Db 171 -----SerLysValSerThrAspAspAlaAlaLysAspValTyrIlePheLeu 187
QY 613 CAAGGATTCCTTAGTGGACTACTAGCTTGGACTTAGGTGCGAGTCTAAGGACTTCAGT 672
Db 188 AspLeuPhePheGluArgPheProHisLeu-----ArgAsnAsnAspPheHis 203
QY 673 CTATGACGAGAGCATGATGAGGCGACTATGTCCTGCATCTTCAATCATTTTACGAG 732
Db 204 IleSerGlyGluSerTyrAlaGlyHisTyrLeuProLysIleAlaHis----- 219
QY 733 CAGATGAGAGAAATGCCAACGGTAGTGTAAATGGTTCAGCTTAATTCAACTCTCTG 792
Db 220 -----GluIleAlaValValHisAlaGluAspSerSerPheAsnLeuSerSerVal 236
QY 793 GGAATTATTAAACGGATCATCGACGAGGCGATCCAGGCCCTTACTACCTGAATTCGT 852
Db 237 LeuIleGlyAsnGlyPheThrAspProLeuThrGlnTyrGlnTyrGluProMetAla 256
QY 853 GTGAACAATACCTACGGTATCAAGGCTGTCAACGAGACCGTCTACAACCTACATGAATT 912
Db 257 CysGlyGluGly---GlyTyrProAlaValLeuGlu----- 267
QY 913 GCCAACCAATGCCAAATGGTTCAGGAT----- 942
Db 268 -----ProGluAspCysLeuAspMetAsnArgAsnLeuProLeuCysLeuSer 283
QY 943 TTGATTTCCACTCGCAACAGACAAACCCAGCCGATAGCTAGTACGCCCTCTCGGCC 1002
Db 284 LeuValAspArgCysTyrLysSerHis-----SerValPheSerCysVal 298
QY 1003 GAAGCCACCAACATGTGAGGAGCAACAATGTAGGGCCATATACGCTTCTGCTGCTGT 1062
Db 299 LeuAlaAspArgTyrCysGluGlnIleThrGly---ValTyrGluLysSerGlyArg 317

```

1063 GGTTGTATGATATTCGG-----CATCATATGATGACCCG-----ACT 1101
318 AsnProTyrAspIleArgSerLysCysGluAlaGluAspAspSerGlyAlaCysTyrGln 337
1102 CCGCAAGTATTACAAATTTCTGGCAAGGACTCTGTCTCATGGAGCGTATCGGCGTC 1161
338 GluGluIleTyrIleSerAspTyrLeuAsnGlnGluValGlnArgAlaLeuGlyThr 357
1162 AACATCAACTACACCCAGTCC---AATAATGACGTCTACTACGTCTTCCAGCAACAGGC 1218
358 AspValSerSerPheGlnGlyCysSerSerAspValGlyIleGlyPheAlaPheThrGly 377
1219 GACTTCTCTGGCCCACTTCATCGAAGACTCGAGAGATCCCTTCTCTCCCGTGGT 1278
378 AspGlyProSerPro---PheHisGlnTyrValAlaGluLeuLeuAspGlnAspIleAsn 396
1279 GTCTCCCTCATCTATGTCGAGCGCGGATTACATCTGCAACTGTTCTGGGGTCCAGCCGTT 1338
397 ValLeuIleTyrAlaGlyAspLysAspTyrIleCysAsnTrpLeuGlyAsnLeuAlaTrp 416
1339 TCCTCTGCTCGAATCACTCCCAAGCGCCGAGTTCGAAAGCGGAGGTACAGCCCTG 1398
417 ThrGluLysLeuGluTrpArgTyrAsnGluGluTyrLysLysGlnValLeuArgThrTrp 436
1399 AAAGTCAACGCGCTCGAG-----TATGGGAAACTCGCGAGTATGTAATTTCTCCTTC 1452
437 LysSerGluGluThrAspGluThrIleGlyGluThrLysSerTyrGlyProLeuThrTyr 456
1453 ACTCGGCTTATGAGCAGCGCCATGAATGCCATCTACCCATCTACCGCCCATCGCTCCCTGCAA 1512
457 LeuArgIleTyrAspAlaGlyHisMetValProHisAspGlnProGluAsnSerLeuGln 476
1513 TTGTTTAAACGGACTATC 1530
477 MetValAsnSerTrpIle 482

RESULT 14
US-08-360-673-4
; Sequence 4, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; NUMBER OF SEQUENCES: 17
; PREPARATION AND USE
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,324
FILING DATE: 23-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/329,892
FILING DATE: 27-OCT-1994
APPLICATION NUMBER: 08/144,704
FILING DATE: 28-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kettleberger, Denise M
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 8648..44USC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/332-5300
TELEFAX: 612/332-9081
TELEX:

INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
08-899-324-33

Alignment Scores:		
Pred. No.:	3.08e-41	Length:
Score:	523.50	Matches:
Percent Similarity:	43.68%	Conservative:
Best Local Similarity:	29.05%	Mismatches:
Query Match:	17.29%	Indels:
DB:	2	Gaps:
		532
		147
		74
		200
		85
		19

US-09-712-338-1 (1-1668) x US-08-899-324-33 (1-532)

QY	97	CCCAGAA	CCCCCACC	CGGGTCAAGACT	-----	CTTACA	ACCGCAACAAT	141	
				:::					
Db	82	ProLysPhe	ProGluAla	IleIleLysThrLysLysAsp	PheValValLysAsnAsp	101			
QY	142	GTCACCA	TCCGGTAC	-----	AAGAA	CCCCGGGCAGAGGC	GCNC	180	
				:::					
Db	102	AlaIleGlu	AsnTyrGlnLeuArgVal	AsnLysIleLysAsp	ProLysIleLeuGlyIle	121			
QY	181	TGCAGACT	ACCCGGGTGC	CAAACTCTACTCT	CGATATGCGACACCTCT	CCCCGAG	---	237	
				:::					
Db	122	-----	AspPro	AsnValThrGlnTyrThr	GlyTyrLeuAspValGluAsp	138			
QY	238	TCCCATC	ACTTCTTCTGGTCT	TCGAA	CCGACAGACATAACCCAGAACT	GCACCTAT	CA	297	
Db	139	LysHisPhe	PheThrThrPheGluSer	ArgAsnAspPro	AlaLysAspPro	ValIle	158		
QY	298	TTCTGGT	TGAATGGTGGCC	TGGAGCGAATCTTTGAT	CGGTCTCTTCGAA	GAGTTGGC	357		
Db	159	LeuTrpLeu	AsnGlyGlyProGlyCys	SerLeuThrGlyLeuPhe	PheGluLeuGly	178			
QY	358	CTTTG	CCATCTCAATTCGACT	TTTGTGACTACATCA	CCCTCACTCGTGGAA	CGAGTC	417		
Db	179	ProSerSer	IleGlyProAspLeu	LysProLleGlyAsnPro	TyrSerTrpAsnSer	Asn	198		
QY	418	TCCAATT	TACTATCTCTG	TCCCCAGCAATGGGAGTCGGCT	TTTTCATATAGTAT	ACGTT	477		
		:::							
Db	199	AlaThrVal	IlePheLeuAspGlnPro	ValAsnValGlyPheSer	TyrSer	215			
QY	478	GATGGG	TCCATTAA	CCCTGTAACTGGGGT	CGTGGAA	AATTTCGAGGAG	AGTT	537	
Db	216	-----							
QY	538	GGCGCG	TACCGCA	ACCATTTGATGCC	ACTCTGATCGAT	ACTACCA	ATCTTGGCG	AGGCC	597

Search completed: November 21, 2003, 17:46:59

Mon Nov 24 13:41:03 2003

us-09-712-338-1.ra1

Page 18

Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 14:49:17; Search time 62 Seconds
(without alignments)
8540.514 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 3027
Sequence: 1 atcggtggtacgaattctt.....ccagtggtgtatgcatag 1668

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Xgapop 10.0	Xgapext 0.5
Xgapop 6.0	Xgapext 7.0
Delop 6.0	Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DBV=xlp
-Q=/cgn2.1/UGPTO.spool.p/US09712338/runat.17112003.170148.10497/app.query.fasta.1.1963
-DB=A Geneseq 19Jun03 -QFMT=fasta -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @C@N 1.1.92 @runat.17112003.170148.10497 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A Geneseq 19Jun03: **
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: **
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: **
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: **
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: **
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: **
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: **
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: **
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: **
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: **
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: **
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: **
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: **
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: **
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: **
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: **
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: **
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: **
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: **
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: **
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: **
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: **
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: **
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: **
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2975	98.3	554	AAW56099	A. oryzae ATCC2038
2	2376	78.5	551	ABR38864	A. niger serine ca
3	980.5	32.4	623	ABR38865	A. niger serine ca
4	947	31.3	554	ABR38859	A. niger serine ca
5	925	30.6	526	ABR38817	A. niger serine ca
6	636	21.0	508	ABG32049	S. cerevisiae BAX-
7	630	20.8	536	ABR38819	A. niger carboxype
8	620	20.5	556	ABR36737	A. niger Bo-1 carb
9	568.5	18.8	491	AAAR48059	A. niger SPAG 2 ca
10	541.5	17.9	550	ABG93281	Sequence of protea
11	419	13.8	481	ABR38845	C. albicans BAX-as
12	383	12.7	482	ABR38845	A. niger carboxype
13	372	12.3	502	ABG30064	Arabidopsis thalia
14	362.5	12.0	476	ABR38843	Arabidopsis thalia
15	361.5	11.9	476	AAW72966	A. niger carboxype
16	361.5	11.9	479	AAW72966	Human membrane or
17	361.5	11.9	479	AAW72966	Human serine carbo
18	361.5	11.9	479	AAW72966	Arabidopsis thalia
19	360.5	11.9	486	AAW72966	Arabidopsis thalia
20	360.5	11.9	486	AAW72966	Amino acid sequenc
21	360.5	11.9	476	AAW72966	Human TANGO 176.
22	360.5	11.9	476	AAW72966	Human PRO polypt
23	360.5	11.9	476	AAW72966	Human hydrophobic
24	360.5	11.9	476	AAW72966	Human PRO223 prote
25	360.5	11.9	476	AAW72966	Novel human secret
26	360.5	11.9	476	AAW72966	Human PRO223 prote
27	360.5	11.9	476	AAW72966	Human PRO polypt
28	360.5	11.9	476	AAW72966	Human secreted/tra
29	360.5	11.9	476	AAW72966	Human secreted/tra
30	360.5	11.9	476	AAW72966	Human secreted/tra
31	360.5	11.9	476	AAW72966	Human secreted/tra
32	360.5	11.9	476	AAW72966	Human secreted/tra
33	360.5	11.9	476	AAW72966	Human secreted/tra
34	360.5	11.9	476	AAW72966	Human secreted/tra
35	360.5	11.9	476	AAW72966	Human secreted/tra
36	360.5	11.9	476	AAW72966	Human PRO polypt
37	360.5	11.9	476	AAW72966	Human secreted/tra
38	360.5	11.9	476	AAW72966	Human secreted/tra
39	360.5	11.9	476	AAW72966	Human secreted/tra
40	360.5	11.9	476	AAW72966	Human secreted/tra
41	360.5	11.9	476	AAW72966	Human secreted/tra
42	360.5	11.9	476	AAW72966	Human secreted/tra
43	360.5	11.9	476	AAW72966	Human secreted/tra
44	360.5	11.9	476	AAW72966	Human secreted/tra
45	360.5	11.9	476	AAW72966	Human secreted/tra

ALIGNMENTS

RESULT 1

AAW56099
ID AAW56099 standard; Protein; 554 AA.

XX
AC AAW56099;

XX
DT 27-AUG-1998 (first entry)

XX
DE A. oryzae ATCC20386 carboxypeptidase I protein.

XX
KW Carboxypeptidase I; flavour improving agent; hydrolyase; proteinaceous; food industry.

XX
OS Aspergillus oryzae.

XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= signal

FT Protein 19..554
 FT /label= carboxypeptidase I
 XX WO9814599-A1.
 XX 09-APR-1998.
 XX 03-OCT-1997; 97WO-US17977.
 XX 27-NOV-1996; 96US-0757534.
 XX 04-OCT-1996; 96US-0726880.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX (NOVO) NOVO-NORDISK AS.
 XX Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;
 XX Klotz A, Mathisen TE, Rey M;
 XX WPI; 1998-240098/21.
 XX N-PSDB; AAV28620.
 XX Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 PT substrates, useful for improving flavour of foods
 PT Claim 1; Fig 3; 82pp; English.
 XX This sequence represents carboxypeptidase I from Aspergillus oryzae.
 CC This polypeptide has an optimal activity in the range of pH 3.0-7.5 at
 CC 25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4, and a
 CC residual activity of at least 65.5% after 30 minutes at pH 4.0 and a
 CC 60 deg. C. It also has the capacity to hydrolyse X from N-CB2-Ala-X where
 CC N-CB2 is N-carboxybenzoxy and X is any amino acid. The carboxypeptidases
 CC can be used for obtaining hydrolysates (which can be enriched in free
 CC glutamic acid or peptide bound glutamic acid residues) from proteinaceous
 CC substrates. The carboxypeptidases can be used in flavour-improving
 CC compositions in the food industry. The products can also be used for the
 CC production of polypeptides free of carboxypeptidase activity.
 XX Sequence 554 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1.09e-259 Length: 554
 Score: 2975.00 Matches: 554
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.28% Indels: 0
 DB: 19 Gaps: 0
 US-09-712-338-1 (1-1668) x AAW56099 (1-554)
 QY 1 ATGGGTGGCTACGAATTTCTCAGTCTACCTTGGTTCAGCCAGTGGGCGCTTCCA 60
 DB 1 MetArgGlyThrGluPheLeuSerValLeuProLeuValAlaAlaSerTrpAlaLeuPro 20
 QY 61 GGAAGTACACGGCGTCCGTGGTAGAAGACAGTACCCAGAACCCACCGGGGTCAAG 120
 DB 21 GlySerThrProAlaSerValGlyArgGlnLeuProLysAsnProThrGlyValLys 40
 QY 121 ACTTTACACCGCAACAATGTCCACCTCGGTACAGGAACCCGGGGAGAGGGCGTC 180
 DB 41 ThrLeuThrThrAlaAsnValThrIleArgGlyLysGluProGlyAlaGluGlyVal 60
 QY 181 TGCAGACTACCCGGGTGTCAAATCTCTCTGATATGTCGACACTCTCCCGAGTCC 240
 DB 61 CysGluThrThrProGlyValLysSerThrGlyThrValAspThrSerProGluSer 80
 QY 241 CATACCTTCTTCTGTTCTTCGAGCCAGACATPACCCAGAACTGCACCTTATCATTG 300
 DB 81 HisThrPheThrTrpPheGluAlaArgHisAsnProGluThrAlaProIleThrLeu 100
 QY 301 TGGTGAATGTGGCCCTCGAAGCGATTCTTTGATCGTCTCTTCGAAAGAGTGGGCCCT 360
 DB 101 TrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluLeuGlyPro 120

QY 361 TGCATGTCAATTCCGACTTTTGATGACTACATCAACCTCCTCCTGCTGGAACGAGGTCTCC 420
 DB CysHisValAsnSerThrPheAspSerTrpIleAsnProHisSerTrpAsnGluValSer 140
 QY 421 AATTACTATTCTCTCCAGCCATTGGGAGTCCGGCTTTTCATATAGTAGACGGTTCAT 480
 DB AsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTrpSerAspThrValAsp 160
 QY 481 GGTCCATTAAACCTGTAACTGGGTGCTCGAAATTCGAGCTTTGAGGAGTTCAGGGC 540
 DB GlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGly 180
 QY 541 CGGTACCCAAACCATTTGATGCTGATCGATACCTACCAATCTTCCGCGCAGAGCCGCT 600
 DB ArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAla 200
 QY 601 TGGGAGATCCTGCAAGGATTCTTGTAGTGGATCTACCTAGCTGGACTCTAGGTGAGTCT 660
 DB TrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSer 220
 QY 661 AAGGACTTCAGTCTATGACGAGAGCTATGAGGGGCTATGGGCTATGGCTCTTCAT 720
 DB LysAspPheSerLeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPheAsn 240
 QY 721 CATTTTACGACGACAGTGAAGATTGCTCAACGCTAGTGTAAATGGTCTTCAGCTTAAT 780
 DB HisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsn 260
 QY 781 TTCAACTCTCTGGGAATTTAAACGGCATCTACGAGGCGATCGAGGCGCTTCTTACTAC 840
 DB PheAsnSerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleGluAlaProTyrTyr 280
 QY 841 CTGTAATTCGTGTGAAACAATPACTACGTATCAGGCTGTCAACGAGACCTCTACAC 900
 DB ProGluPheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsn 300
 QY 901 TACATGAAGTTTCCCAACCAATGCCAATGGTTCAGGATTTGATTTCCACTGCAAA 960
 DB TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys 320
 QY 961 CAGACAAACCGCACCGCATTTAGTCTAGCTACGCTCTGCGCGCAAGCCACCAACATGTGC 1020
 DB GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys 340
 QY 1021 AGGACAATTTGAGGGCCATCTACTAGCCTTTGCTGGTCTGCTGTGTATGATATTCGG 1080
 DB ArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArg 360
 QY 1081 CATCCATATGATGACCGCATCTCCGCAAGTTATTACAACAAATTTCTGGCAAGGACTCT 1140
 DB HisProTyrAspAspProThrProProSerTyrTyrAsnLysPheLeuAlaLysAspSer 380
 QY 1141 GTCATGACGCTATCCGCGTCACATCACTACACCCAGTCCCAATATAGCTCTACTAC 1200
 DB ValMetAspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyr 400
 QY 1201 GCTTTCAGAAACAGCGACTTGTCTGCCCAACTTCATCGAAGACCTCGAGGAGATC 1260
 DB AlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluGluIle 420
 QY 1261 CTTGCTCTCCCGTGGCTGTCTCCCTCATCTATGGCGACCGCCGATTCATCTGCAACTGG 1320
 DB LeuAlaLeuProValAlaGValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrp 440
 QY 1321 TTCGGCGGTGAGCGCTTTCCTCGTGGCAACTACTCCCAAGCGCGCGATTCGCGAGC 1380
 DB PheGlyGlyGlnAlaValSerLeuAlaAsnTyrSerGlnAlaAlaGlnPheArgSer 460
 QY 1381 GCAGGTACACGCGCTTGAAGTCAACGGCTCGAGTATGGGAAATCTCGCGAGTATGGT 1440
 DB AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly 480

```

QY 1441 AATTCTCCTTCACTCGGCTGATAGGAGGCGCATGAAGTCCCATACTACCGCCCATC 1500
Db |||||
QY 481 AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProile 500
Db |||||
QY 1501 GCCTCCCTGCAATTTAAACCGGACTATCTCGTTGGGATATCGAGAGGGCCAGAAG 1560
Db |||||
QY 501 AlasSerLeuGlnLeuPheAsnArgThrIlePheGlyTrpAspIleAlaGluGlyGlnLys 520
Db |||||
QY 1561 AAGATCTGGCCCACTACAAGACGAATGGAACGGCTACAGCTAGCATACATACACATCGCTCC 1620
Db |||||
QY 521 LysIleTrpProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer 540
Db |||||
QY 1621 GTCCGCTGCTACGGCTACAGCATGTCCAGTTCTGGTATG 1662
Db |||||
QY 541 ValProLeuProThrAlaThrSerMetSerSerValGlyMet 554
Db |||||

RESULT 2
ABR38864
ID ABR38864 standard; Protein; 551 AA.
AC ABR38864;
XX
XX
XX 24-APR-2003 (first entry)
XX
XX
XX A. niger serine carboxypeptidase polypeptide #3.
XX
XX Protease; fungal infection; aspergilliosis; food; tanning; detergent;
KW protein solubility; viscosity; taste; texture; nutritional value;
KW EC3.4.16.6.
XX
XX Aspergillus niger.
XX
XX WO200268623-A2.
XX
XX 06-SEP-2002.
XX
XX 22-FEB-2002; 2002WO-EP01984.
XX
XX 23-FEB-2001; 2001EP-0200657.
PR 23-FEB-2001; 2001EP-0200658.
PR 23-FEB-2001; 2001EP-0200660.
PR 26-FEB-2001; 2001EP-0200706.
PR 26-FEB-2001; 2001EP-0200707.
PR 26-FEB-2001; 2001EP-0200708.
PR 26-FEB-2001; 2001EP-0200719.
PR 28-MAR-2001; 2001EP-0000075.
PR 28-MAR-2001; 2001EP-0000078.
PR 28-MAR-2001; 2001EP-0000080.
PR 28-MAR-2001; 2001EP-0000087.
PR 28-MAR-2001; 2001EP-0000088.
PR 21-MAY-2001; 2001EP-0000156.
PR 21-MAY-2001; 2001EP-0000159.
PR 21-MAY-2001; 2001EP-0000160.
PR 21-MAY-2001; 2001EP-0000162.
PR 21-MAY-2001; 2001EP-0000165.
PR 21-MAY-2001; 2001EP-0000166.
PR 21-MAY-2001; 2001EP-0000168.
PR 21-JUN-2001; 2001EP-0000240.
PR 21-JUN-2001; 2001EP-0000242.
PR 21-JUN-2001; 2001EP-0000244.
PR 21-JUN-2001; 2001EP-0000246.
PR 12-JUL-2001; 2001EP-0000280.
PR 12-JUL-2001; 2001EP-0000285.
PR 30-JUL-2001; 2001EP-0000323.
PR 30-JUL-2001; 2001EP-0000327.
PR 02-AUG-2001; 2001EP-0000341.
PR 02-AUG-2001; 2001EP-0000342.
PR 02-AUG-2001; 2001EP-0000343.
PR 02-AUG-2001; 2001EP-0000344.
PR 09-AUG-2001; 2001EP-0000357.
PR 16-AUG-2001; 2001EP-0000374.
PR 16-AUG-2001; 2001EP-0000377.
PR 20-SEP-2001; 2001EP-0000478.

```

```

PR 20-SEP-2001; 2001EP-0000483.
PR 22-OCT-2001; 2001EP-0000552.
PR 22-OCT-2001; 2001EP-0000553.
PR 22-OCT-2001; 2001EP-0000554.
PR 22-OCT-2001; 2001EP-0000555.
PR 22-OCT-2001; 2001EP-0000556.
PR 22-OCT-2001; 2001EP-0000557.
PR 22-OCT-2001; 2001EP-0000558.
PR 15-NOV-2001; 2001EP-0004464.
PR 21-DEC-2001; 2001EP-0005117.
XX
XX (STAM ) DSM NV.
XX
XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
PI Stiebler J, Albang R;
XX
XX WPI; 2002-723203/78.
DR
DR P-PSDB; ABZ78231, ABZ78288.
XX
XX Novel isolated protease polypeptide useful in laboratory, clinical,
PT pharmaceutical, chemical, diagnostic, personal care and industrial
PT applications.
XX
XX Claim 13; Page 379-382; 394pp; English.
XX
XX The invention relates to a novel isolated protease polypeptide. A
CC polypeptide or polynucleotide of the invention is useful for diagnosing a
CC fungal infection such as aspergilliosis, or as a query sequence to perform
CC a search against public databases. A polypeptide of the invention is
CC useful in a selected number of industrial or pharmaceutical processes, in
CC laboratory or clinical processes, in food industry (baking, brewing,
CC cheese manufacturing, meat tenderising), in tanning industry and in the
CC manufacture of biological detergents. A polypeptide may also be useful
CC for improving protein solubility, extraction yields, viscosity or taste,
CC texture, nutritional value, minimising of antigenicity or
CC anti-nutritional factors, colour or functionality as well as processing
CC aspects like filterability of the proteinaceous raw material. The
CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
CC the invention.
XX
XX Sequence 551 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1.39e-205 Length: 551
XX Score: 2376.00 Matches: 428
XX Percent Similarity: 87.96% Conservative: 54
XX Best Local Similarity: 78.10% Mismatches: 66
XX Query Match: 78.49% Indels: 0
XX DB: 23 Gaps: 0
XX
XX US-09-712-338-1 (1-1669) x ABR38864 (1-551)
QY 1 ATGCGTGGCTACGAATTTCTCTCAGTGTACCCCTGGTTGACGCCAGTTGGCCCTTCCA 60
Db |||||
QY 1 MetArgGlySerArgLeuValLeuLeuLeuProLeuAlaLeuSerCysAlaMetPro 20
Db |||||
QY 61 GGAAGTACACCGGCGTCGTCGTTAGAGACAGCTACCCAAAGAACCCACCGGGGTCAAG 120
Db |||||
QY 21 GluAsnGluTrpSerSerThrIleArgArgGlnLeuProLysAlaSerThrGlyVallys 40
Db |||||
QY 121 ACTCTTACACCGCAACAATGTACCATCCGGTACAAAGGAACCCGGGCGAGGCGTC 180
Db |||||
QY 41 SerIleLysThrProAsnAsnValThrIleArgTyrLysGluProGlyThrGluGlyIle 60
Db |||||
QY 181 TGCAGAGCTACCCGGGTGTCAAATCTCTACTCTGGATATCGACACCTCTCCCGAGTCC 240
Db |||||
QY 61 CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspLeuSerProGluSer 80
Db |||||
QY 241 CATACCTTCTTCTGTTCTTTCGAAGCCAGACATAACCCAGAAACTGCACCTATCATGTC 300
Db |||||
QY 81 HisThrPhePheTrpPheGluSerArgArgAspProGluAsnAspProValThrLeu 100
Db |||||

```

QY	301	TGGTTGAATGGTGGCCCTGGAAGCGATTCTTTTGATCGCTCTCTTCGAAGAGTTCGGCCCT	360	Db	461	AlaGlyTyrThrProMetThrValAspGlyValGluTyrGlyGluThrArgGluTyrGly	480
Db	101	TrpLeuasnGlyGlyProGlySerAspSerLeuileuileGlyLeuPheGluGluLeuGlyPro	120	QY	1441	AATTTTCCTTCCTCACTCGCGTCTATGAGCAGGCCCATGAAGTCCCATACTACAGCCCATC	1500
QY	361	TGCATGTCAAATCGACTTTTGATGACATACATACACCTCACTCGTGGAAAGAGTCTCC	420	Db	481	AsnPheSerPheThrArgValTyrGlnAlaGlyHisGluValProTyrTyrGlnProile	500
Db	121	CysHisIleThrProGluTyrGluSerIleIleAsnGlnTyrSerTyrPasnGluValThr	140	QY	1501	GCCTCCCTGCAATTTTAAACCGGACTATCTTCGTTGGGATATCGAGAGGCCAGAG	1560
QY	421	AATTACTATTCCTGTCGCCACCATTTGGAGTTCGGCTTTTCATATAGTATACGTTGAT	480	Db	501	AlaAlaLeuGlnLeuPheAsnArgThrLeuPheGlyTyrAspIleAlaAlaGlyThrThr	520
Db	141	AsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerGluThrGluAla	160	QY	1561	AAGATCTCGCCCGCTACAGACCAATGCAACGGCTACAGCTACGCATACACAGTCTGCC	1620
QY	481	GGGTCCATTAACCTGTAACTGGGGTCTCGAAAATTCGAGCTTTCGAGGAGTTCAGGCG	540	Db	521	GlnIleTyrProGluTyrSerThrAsnGlyThrSerGlnAlaThrHisThrGluSerPhe	540
Db	161	GlySerLeuAsnProPheThrGlyAlaValGluAsnAlaSerPheAlaGlyValGlnGly	180	QY	1621	GTGCGCTCGCTACGCTACCGTACCAGC	1644
QY	541	CGGTACCAACCAATGATGCGCACTCTGATCGATACCTACCAATCTTCGCCGAGAGCCGCT	600	Db	541	ValProLeuSerThrAlaSerSer	548
Db	181	ArgTyrProValIleAspAlaThrIleIleAspThrThrAspIleAlaAlaArgAlaThr	200	RESULT 3			
QY	601	TGGGAGATCTCGCAAGGATTCCTTAGTGACTACCTAGCTTGGACTCTAGGCTGCGAGTCT	660	ID	ABR38865	standard; Protein; 623 AA.	
Db	201	TrpGluValLeuGlnGlyPheLeuSerGlyLeuSerGlnLeuAspSerGluValIysSer	220	XX	ABR38865;		
QY	661	AAGGACTTCAGTCTATGACGAGAGCTATGAGGCGCATGCTGCTCGCTCATCTTCTAAT	720	XX	24-APR-2003	(first entry)	
Db	221	LysGluPheAsnLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPheAsn	240	DE	A. niger serine carboxypeptidase polypeptide #4.		
QY	721	CATTTTACGACAGATGAGAGATTCACAGGTAGTTAATGGTCTTCAGCTTAAT	780	KW	Protease; fungal infection; aspergillosis; food; tanning; detergent;		
Db	241	HisPheTyrGluGlnAsnSerLysIleAlaSerGlyGluValAsnGlyValGlnLeuAsn	260	KW	protein solubility; viscosity; taste; texture; nutritional value;		
QY	781	TTCAACTCTCTGGGAATTAATTAACGCATCATCGACGCGCATCCAGGCCCTTACTAC	840	XX	EC3.4.16.6.		
Db	261	PheAsnSerLeuGlyIleIleAsnGlyIleIleAspAlaAlaIleGlnAlaAspTyrTyr	280	OS	Aspergillus niger.		
QY	841	CCTGAATTCGTGTGAACAATACCTACGGTATCAAGCTGTCAACGAGACCGTCTCAAC	900	XX	WO200268623-A2.		
Db	281	AlaAspPheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnAspThrValTyrAsn	300	PD	06-SEP-2002.		
QY	901	TACATGAAGTTTCCCAACCAATGCAATGGTTCGAGGATTTGATTCACCTCGAAA	960	XX	22-FEB-2002; 2002WO-EP01984.		
Db	301	TyrMetLysPheAlaAsnThrMetProAsnGlyCysGlnAspGlnValAlaSerCysLys	320	XX	23-FEB-2001; 2001EP-0200657.		
QY	961	CAGACAAACCGACCGCATTAGCTACTAGCCCTCTGGCCGAGACCGACCAACATGTGC	1020	PR	23-FEB-2001; 2001EP-0200658.		
Db	321	LeuThrAsnArgThrSerLeuSerAspTyrAlaIleCysThrGluAlaAlaAsnMetCys	340	PR	26-FEB-2001; 2001EP-0200660.		
QY	1021	AGGCAATGTTGAGGGCCCATACTACGCCCTTTCGCTGGTGGTGTGTATGATATCGG	1080	PR	26-FEB-2001; 2001EP-0200706.		
Db	341	ArgAspAsnValGluGlyProTyrTyrGlnPheGlyArgGlyValTyrAspIleArg	360	PR	26-FEB-2001; 2001EP-0200708.		
QY	1081	CATCCATATGATACCGGCTCCGCAAGTATTACAAATTTCTGGCAAGGACTCT	1140	PR	26-FEB-2001; 2001EP-0200719.		
Db	361	HisProTyrAsnAspProThrProProSerTyrPheValAspTyrIleuLysAspSer	380	PR	28-MAR-2001; 2001EP-0000075.		
QY	1141	GTATGACGCTATCGCGCTCAACATCACTACACCGGCTCCATATGACGCTACTAC	1200	PR	28-MAR-2001; 2001EP-0000078.		
Db	381	ValMetAspAlaIleGlyValAspIleAsnTyrThrGluSerSerGlyGluValTyrTyr	400	PR	28-MAR-2001; 2001EP-0000080.		
QY	1201	GCTTTCGAGAAACAGCGCACTTGTGTGCGCACTTTCATCGAAGACCTCGAGGATC	1260	PR	28-MAR-2001; 2001EP-0000087.		
Db	401	AlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluIle	420	PR	28-MAR-2001; 2001EP-0000088.		
QY	1261	CTTCTCTCCCGGCGTGTCTCTCTATCTATCGGACCGCGGATTCATCTGCAACTGG	1320	PR	21-MAY-2001; 2001EP-0000156.		
Db	421	LeuGlnLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrp	440	PR	21-MAY-2001; 2001EP-0000159.		
QY	1321	TTGGCGGCTGAGGCGGTTTCTCTCGTCGCACTACTCCCAAGCGCGGCTTCGGAAGC	1380	PR	21-MAY-2001; 2001EP-0000160.		
Db	441	PheGlyGlyGlnAlaIleSerLeuAlaValAsnTyrProHisAlaAlaGlnPheArgAla	460	PR	21-MAY-2001; 2001EP-0000162.		
QY	1381	GCAGGGTACAGCCCTCGAAAGTCAACGGCGTCTGAGTATGGGGAACCTCGCGAGTATGGT	1440	PR	21-MAY-2001; 2001EP-0000165.		
Db				PR	21-MAY-2001; 2001EP-0000166.		
				PR	21-MAY-2001; 2001EP-0000168.		
				PR	21-JUN-2001; 2001EP-0000240.		
				PR	21-JUN-2001; 2001EP-0000242.		
				PR	21-JUN-2001; 2001EP-0000246.		
				PR	12-JUL-2001; 2001EP-0000280.		
				PR	12-JUL-2001; 2001EP-0000285.		
				PR	30-JUL-2001; 2001EP-0000323.		
				PR	30-JUL-2001; 2001EP-0000327.		
				PR	02-AUG-2001; 2001EP-0000341.		
				PR	02-AUG-2001; 2001EP-0000342.		
				PR	02-AUG-2001; 2001EP-0000343.		
				PR	02-AUG-2001; 2001EP-0000344.		
				PR	09-AUG-2001; 2001EP-0000357.		

Db 459 ArgGlyGlyGluAspValSerLeuValGluTyrGluAspAlaGluLysPheArgala 478
Qy 1381 GCAGGTACACGCCCTCAAGTCAACGGCGTCGAGTAT-----GGGAAACTCGCGAG 1434
Db 479 AlaGlyTyrAlaGluValGlnThrLysSer---SeryrValGlyGlyLeuValArgGln 497
Qy 1435 TATGTAATTTCTCTCTCACTCGGCTCTATGAGGAGCGCCATGAGTCCCATACACAG 1494
Db 498 TyrGlyAsnPheSerPheThrArgValPheGlnAlaGlyHisGluValProPheTyrGln 517
Qy 1495 CCCATCGGCTCCCTGCAATGTTAAACCGGACTATCTTCGGTTGGATATGCGAGAGGC 1554
Db 518 ProGluThrAlaTyrGluLeuPheAsnArgAlaGlnPheAsnTrpAspIleAlaThrGly 537
Qy 1555 CAGAAGAGATCGG-----CCACGCTACAGACGAATGGAAACGCTACAGCTACGCAT 1608
Db 538 GlyIleSerLeuGluGlnAsnGlnSerTyrGlyThrGluGlyProSerSerThrTrpHis 557
Qy 1609 ACACAGTCTCGTGGCGG 1626
Db 558 IleLysAsnGluValPro 563
RESULT 4
ABR38859
ID ABR38859 standard; Protein; 554 AA.
XX AC ABR38859;
DT
DT 24-APR-2003 (first entry)
XX
DE A. niger serine carboxypeptidase polypeptide #2.
XX
XX
KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
KW protein solubility; viscosity; taste; texture; nutritional value;
KW EC3.4.16.6.
OS Aspergillus niger.
XX
PN WC200268623-A2.
XX
XX
XX 06-SEP-2002.
XX
XX 22-FEB-2002; 2002WO-EP01984.
XX
XX 23-FEB-2001; 2001EP-0200657.
XX 23-FEB-2001; 2001EP-0200658.
XX 23-FEB-2001; 2001EP-0200660.
XX 26-FEB-2001; 2001EP-0200706.
XX 26-FEB-2001; 2001EP-0200707.
XX 26-FEB-2001; 2001EP-0200708.
XX 26-FEB-2001; 2001EP-0200719.
XX 28-MAR-2001; 2001EP-0000075.
XX 28-MAR-2001; 2001EP-0000078.
XX 28-MAR-2001; 2001EP-0000080.
XX 28-MAR-2001; 2001EP-0000087.
XX 28-MAR-2001; 2001EP-0000088.
XX 21-MAY-2001; 2001EP-0000156.
XX 21-MAY-2001; 2001EP-0000159.
XX 21-MAY-2001; 2001EP-0000160.
XX 21-MAY-2001; 2001EP-0000162.
XX 21-MAY-2001; 2001EP-0000165.
XX 21-MAY-2001; 2001EP-0000166.
XX 21-MAY-2001; 2001EP-0000168.
XX 21-MAY-2001; 2001EP-0000169.
XX 21-JUN-2001; 2001EP-0000240.
XX 21-JUN-2001; 2001EP-0000242.
XX 21-JUN-2001; 2001EP-0000244.
XX 21-JUN-2001; 2001EP-0000246.
XX 12-JUL-2001; 2001EP-0000280.
XX 12-JUL-2001; 2001EP-0000285.
XX 30-JUL-2001; 2001EP-0000323.
XX 30-JUL-2001; 2001EP-0000327.
XX 02-AUG-2001; 2001EP-0000341.
XX 02-AUG-2001; 2001EP-0000342.

PR 02-AUG-2001; 2001EP-0000343.
PR 02-AUG-2001; 2001EP-0000344.
PR 09-AUG-2001; 2001EP-0000357.
PR 16-AUG-2001; 2001EP-0000374.
PR 16-AUG-2001; 2001EP-0000377.
PR 20-SEP-2001; 2001EP-0000478.
PR 20-SEP-2001; 2001EP-0000483.
PR 22-OCT-2001; 2001EP-0000552.
PR 22-OCT-2001; 2001EP-0000553.
PR 22-OCT-2001; 2001EP-0000554.
PR 22-OCT-2001; 2001EP-0000556.
PR 22-OCT-2001; 2001EP-0000557.
PR 22-OCT-2001; 2001EP-0000558.
PR 15-NOV-2001; 2001EP-0004464.
PR 21-DEC-2001; 2001EP-0005117.
XX (STAM) DSM NV.
XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
PI Klingbauer S, Wagner C, Fritz A, Von Gustadt W, Heinrich O;
PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
PI Stiebeler J, Albarg R;
XX
XX WPI: 2002-723203/78.
DR P-PSDB; ABZ78226, ABZ78283.
XX
PT Novel isolated protease polypeptide useful in laboratory, clinical,
PT pharmaceutical, chemical, diagnostic, personal care and industrial
PT applications -
XX
PS Claim 13; Page 366-369; 394pp; English.
XX
CC The invention relates to a novel isolated protease polypeptide. A
CC polypeptide or polynucleotide of the invention is useful for diagnosing a
CC fungal infection such as aspergillosis, or as a query sequence to perform
CC a search against public databases. A polypeptide of the invention is
CC useful in a selected number of industrial or pharmaceutical processes, in
CC laboratory or clinical processes, in food industry (baking, brewing,
CC cheese manufacturing, meat tenderising), in tanning industry and in the
CC manufacture of biological detergents. A polypeptide may also be useful
CC for improving protein solubility, extraction yields, viscosity or taste,
CC texture, nutritional value, minimising of antigenicity or
CC anti-nutritional factors, colour or functionality as well as processing
CC aspects like filterability of the proteinaceous raw material. The
CC sequences shown in ABR38313-ABR38889 represent the A. niger proteases of
CC the invention.
XX
SQ Sequence 554 AA;
Alignment Scores:
Pred. No.: 1.67e-76 Length: 554
Score: 947.00 Matches: 215
Percent Similarity: 51.15% Conservatives: 73
Best Local Similarity: 38.19% Mismatches: 176
Query Match: 31.29% Indels: 99
Gaps: 14
DB:
US-09-712-338-1 (1-1668) x ABR38859 (1-554)
Qy 91 CAGCTACCCCAAGAACCCACCGGGGTCAAGACTTTTACACCGCA-----ACAATGTC 144
Db 17 GlnPheProGluProGluGlyIleThrValLeuLysSerLysLeuHisGluAsnVal 36
Qy 145 ACCATCGGTACAGGACCGGGGCGAGAGCTCTCCGAGTACCTACCTCTTCTGTTCTTCGAA 204
Db 37 ThrIleSerPheLysGluPro-----GlyIleCysGluThrThrProGlyValArg 53
Qy 205 TCCTACTCTGATATGTCGACACACCTCTCCCGAGTCCCATACCTCTTCTGTTCTTCGAA 264
Db 54 SeryrSerGlyTyrValHisLeuProAlaSerThrSerPhePheThrPheGlu 73
Qy 265 GCCAGACATAACCCAGAAACTGCACCTATCATCTGTTGTTGATGGTGGCCCTGGAAGC 324

Db 74 AlaArgLysAspProSerAsnAlaProLeuAlaIleTrpLeuAsnGlyGlyProGlyGly 93
 QY 325 GATTCTTTGATCGGTCTCTTCAAGAGTGGCGCTTGCATGTCAATTGCG---ACTTTT 381
 Db 94 SerSerLeuMetGlyLeuLeuGluLeuGlyProCysSerIleAlaSerAspSerLys 113
 QY 382 GATGACTACATCAACCTCTACTCGTGGACAGAGTCTCCAATTTACTATTCTCTGCCAG 441
 Db 114 ThrThrValLeuAsnProTrpSerTrpAsnAsnGluValAsnLeuLeuPheLeuAspGln 133
 QY 442 CCATTGGAGTCCGCTTTTCATATAGTCATACGGTGTGATGGTCCATTAACCCCTGTAACT 501
 Db 134 ProThrGlnValGlyPheSerTrpAspValProThrAsnGlyThrLeuThr----- 150
 QY 502 GGGTCTGTCGAAAATTCGAGCTTTGCAGGAGTTTCAGGGCCGTTACCCCAACCAATGATGCC 561
 Db 150 ----- 150
 QY 562 ACTCTGATCGTACTACCAATCTTGGCGAGAGCGCTTGGGAGATCTTSCAAGGATTC 621
 Db 151 -----AlaAsnGlyThrAlaPheAlaAlaHisAlaLeuTrpHisPheAlaGlnThrTrp 168
 QY 622 CTTAGTGGACTTACCTAGCTTG-----GACTCTAGGGTGCAGTCTAAGGACTTCAGT 672
 Db 169 PhePheGluPheProHisTyrLysProAsnAspArgVal-----Ser 183
 QY 673 CTATGCGGAGAGTATGAGGCGCACTATGCTCTGCAATCTTCAATCTTTTACGAG 732
 Db 184 LeuTrpAlaGluSerTyrGlyHisTyrGlyProGlyIlePheArgPheGlnGln 203
 QY 733 CAGATCAGAGATTTGCAACGGTAGTGT---AATGGTGTTCAG---CTTAATTTCAAC 786
 Db 204 GlnAsnAspLysIleAlaGluGlyThrAlaGluAspGlyAlaGlnTyrLeuHisLeuAsp 223
 QY 787 TCTCTGGGAATTATTAACGGCATCATCGACGAGCGATCCAGGCCCTTACTACCTCGAA 846
 Db 224 ThrLeuGlyIleValAsnGlyLeuMetAspMetValIleGlnGluGluAlaTyrIleThr 243
 QY 847 TTCGCTGTGAACAATACCTACGGTATCAAGCTGTCAACGAGACCGTCTACACATCATG 906
 Db 244 TrpPro-----TyrAsnAsnVal 249
 QY 907 AGTTTTCGCAACCAATGCCAAATGGTTGCCAGATTGATTTCCACCTGCAACACAGACA 966
 Db 250 ArgLeuAlaPro-----SerSerPhe 256
 QY 967 AACCGCACCGCATTAGTCTACTACGCTCTGCGCCGACGCCACCAACATGTCCAGGAC 1026
 Db 257 AsnSerArgGlyPheArgAspGlnAlaLeuAlaCysGluAlaLeuTyrGluArgAsp 276
 QY 1027 -----AATGTT----- 1032
 Db 277 SerGlyLeuProHisSerGlyLysAsnIleSerGluIleCysGlyLeuAlaLeuGlu 296
 QY 1033 -----GAGGGGCA-----TACTACGCTTTTGTGGTGTGGTGTGTATGATATTCGG 1080
 Db 297 TrpGlyAspGlyProIleThrTyrTyrHisThrPheAsnArgGlyTrpTyrAspIleAla 316
 QY 1081 CATCCATATGATACCGGACTCCGCCAAGTATTACAAATTTCTGGCAAGGACTCT 1140
 Db 317 HisProGlyAsnAspProPheProAlaLysHisMetLeuGlyTyrLeuThrGlnGluSer 336
 QY 1141 GTCATGAGCGTATCGCGCTCAACATCACTACACCGAGTCCCAATATGACGCTACTAC 1200
 Db 337 ValLeuAlaAlaLeuGlyValProValAsnPheThrSerSerSerAlaValAlaThr 356
 QY 1201 GCTTTCAGCAACACGGGACTTTGCTGGCCCAACTTCAATCGAAGACCTCCAGGAGATC 1260
 Db 357 GlnPheIleLysThrPheAspIleValHisGlyGlyPheLeuAspAlaIleGlyTyrLeu 376
 QY 1261 CTGTCTCTCCCGTGGTCTCCCTCATCTATGGCAGCGCCGATTAATCTGCAACTGG 1320
 Db 377 LeuAspSerGlyValLysValHisMetMetTyrGlyAspArgAspTyrAlaCysAsnTrp 396

QY 1321 TTCGGCGTACGGCGGTTTCCTCGTCGCGAATCTATCCCAAGCGCCGATTCGGAAGC 1380
 Db 397 ValGlyGlyGluLysAlaSerLeuAlaValProTyrSerArgIleThrGluPheAlaAsp 416
 QY 1381 GCAGGTACACGGCCCTGAAAGTCAACGGCGTCGAGTATGGGAAACTCGCGAGTATCGT 1440
 Db 417 ThrGlyTyrSerProLeuLeuThrProAspGlyIleSerGlyMetThrArgGlnLeuGly 436
 QY 1441 AATTCTCTCTACTCGCTCTATAGGCGAGCCATGAAGTCCCATACTACAGCCCATC 1500
 Db 437 AsnTyrSerPheThrArgValPheGlnAlaGlyHisGluValProSerTyrGlnProVal 456
 QY 1501 GCCTCCCTGCAATTTTAAACCGGACTATCTTCGTTGGGATATCGCAGAGGGCCAGAAG 1560
 Db 457 AlaAlaTyrGluIlePheMetArgAlaThrPheAsnLysAspIleProThrGlyLeuLeu 476
 QY 1561 AAGATCTGGCCGCTCAAGACGAATGAACGCTACAGCTACGCTACACATACACATCGTCC 1620
 Db 477 AlaValAspAspGluPheGlnSerValGlyProLysAspThrTrpHisIleLysAsnIle 496
 QY 1621 GTGCGCTGCTACGGCTACCGCATCGCATGTCCAGTGTTC----- 1657
 Db 497 ProProIle-MetProLysProGlnCysTyrValLeuSerProGlyThrCysThrProGly 516
 QY 1658 -GTATGG 1663
 Db 516 uValTrp 518
 RESULT 5
 ABR38817
 ID ABR38817 standard; Protein; 526 AA.
 XX AC ABR38817;
 XX DT 24-APR-2003 (first entry)
 XX DE A. niger serine carboxypeptidase polypeptide #1.
 KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.6.
 XX OS Aspergillus niger.
 XX PN WO200268623-A2.
 XX PD 06-SEP-2002.
 XX PF 22-FEB-2002; 2002WO-EP01984.
 XX PR 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0000075.
 PR 28-MAR-2001; 2001EP-0000078.
 PR 28-MAR-2001; 2001EP-0000080.
 PR 28-MAR-2001; 2001EP-0000087.
 PR 28-MAR-2001; 2001EP-0000088.
 PR 21-MAY-2001; 2001EP-0000156.
 PR 21-MAY-2001; 2001EP-0000159.
 PR 21-MAY-2001; 2001EP-0000160.
 PR 21-MAY-2001; 2001EP-0000162.
 PR 21-MAY-2001; 2001EP-0000165.
 PR 21-MAY-2001; 2001EP-0000166.
 PR 21-MAY-2001; 2001EP-0000168.
 PR 21-JUN-2001; 2001EP-0000240.
 PR 21-JUN-2001; 2001EP-0000242.
 PR 21-JUN-2001; 2001EP-0000244.

PR 21-JUN-2001; 2001EP-0000246.
 PR 12-JUL-2001; 2001EP-0000280.
 PR 12-JUL-2001; 2001EP-0000285.
 PR 30-JUL-2001; 2001EP-0000323.
 PR 30-JUL-2001; 2001EP-0000327.
 PR 02-AUG-2001; 2001EP-0000341.
 PR 02-AUG-2001; 2001EP-0000342.
 PR 02-AUG-2001; 2001EP-0000343.
 PR 02-AUG-2001; 2001EP-0000344.
 PR 02-AUG-2001; 2001EP-0000357.
 PR 16-AUG-2001; 2001EP-0000374.
 PR 16-AUG-2001; 2001EP-0000377.
 PR 20-SEP-2001; 2001EP-0000478.
 PR 20-SEP-2001; 2001EP-0000483.
 PR 22-OCT-2001; 2001EP-0000552.
 PR 22-OCT-2001; 2001EP-0000553.
 PR 22-OCT-2001; 2001EP-0000554.
 PR 22-OCT-2001; 2001EP-0000556.
 PR 22-OCT-2001; 2001EP-0000557.
 PR 22-OCT-2001; 2001EP-0000558.
 PR 15-NOV-2001; 2001EP-0000446.
 PR 21-DEC-2001; 2001EP-00005117.
 XX (STAM) DSM NV.

PI Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 PI Klugbauer S, Wagner C, Fritz A, Von Gusted W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebeler J, Albang R;

XX WPI; 2002-723203/78.
 DR P-PSDB; AB278184, AB278241.

XX Novel isolated protease polypeptide useful in laboratory, clinical,
 PT pharmaceutical, chemical, diagnostic, personal care and industrial
 PT applications

XX Claim 13; Page 271-274; 394pp; English.

XX The invention relates to a novel isolated protease polypeptide. A
 CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 CC fungal infection such as aspergillosis, or as a query sequence to perform
 CC a search against public databases. A polypeptide of the invention is
 CC useful in a selected number of industrial or pharmaceutical processes, in
 CC laboratory or clinical processes, in food industry (baking, brewing,
 CC cheese manufacture, meat tenderising), in tanning industry and in the
 CC manufacture of biological detergents. A polypeptide may also be useful
 CC for improving protein solubility, extraction yields, viscosity or taste,
 CC texture, nutritional value, minimising of antigenicity or
 CC anti-nutritional factors, colour or functionality as well as processing
 CC aspects like filterability of the proteinaceous raw material. The
 CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
 CC the invention.

XX Sequence 526 AA;

Alignment Scores:
 Pred. No.: 1.59e-74 Length: 526
 Score: 925.00 Matches: 221
 Percent Similarity: 51.46% Conservative: 78
 Best Local Similarity: 38.04% Mismatches: 192
 Query Match: 30.56% Indels: 90
 DB: 23 Gaps: 15

US-09-712-338-1 (1-1668) x ABR38817 (1-526)

QY 10 TACGAATTTCTCAGTCCTACCTTGGTTCAGCCAGTGGCGCTTCCAGGAGTACA 69
 Db 2 TyrTyrSerLeuTrpValAlaAlaLeuValAla-----AlaLeuPro----- 15
 QY 70 CCGGCGTCCGTCGGTAGAAGACAGCTACCCCAAGAACCCCGGGGTCAAGACTCTTACA 129
 Db 16 -----ValSerArgAlaGlnPheValAlaProProThrAspLeuIleProThrlys 32

QY 130 ACCGCAACAATGTCACCATCGGTACAGAACCCGGGGCAGAGGGCGTCTGCAGACT 189
 Db 33 GlyTyrLeuAspIleProValArgTyrLysGlnVal---ProThrGlyIleCysGluThr 51
 QY 190 ACCCGGGTGTCAAATCTCTCTGGATATGCGACACCTCTCCGAGTCCCATACCTTC 249
 Db 52 AspProSerValLysSerPheSerGlyTyrValAspValAlaGluHisGluHisIlePhe 71
 QY 250 TTCTGTTCTTTCGAAGCCAGACAT---AACCCAGAAACTGCACCTATCATCATCTGGTGG 306
 Db 72 PheTrpPhePheGluAlaArgAsnGlnAspProThrGluAlaProLeuThrValTrpIle 91
 QY 307 AATGGT-----GGCCTCGAAGCGATCTTTCATCGGTCTCTTCCGAAGAG 351
 Db 92 AsnGlyGlyMetSerAspProGlyProGlySerSerMetIleGlyLeuPheGlnGlu 111
 QY 352 TTGGGCGCTTGCATGCAATTCGACTTTTGTGACTACATCAACCTCCTCGTGGAAC 411
 Db 112 HisGlyProCysGlyIleAspAlaAsnGlySerValTyrAsnAsnProTyrSerTrpAsn 131
 QY 412 GAGGTCTCAAATTTACTATCTCTGCCAGCCATCGGAGTCGGCTTTTCATATAGTAT 471
 Db 132 AsnAlaSerAsnMetLeuTyrIleAspGlnProValGlnThrGlyPheSerTyrSer--- 150
 QY 472 ACGTGTGAGGGTCCATTAACTGTAACCTGTAACTGGGGTCTCGAAATTCGAGCTTGCAGGA 531
 Db 151 -----IleProValProGlyTyrValAspSerSerThrAspAsnGly 164
 QY 532 GTTCAGGGCGGTACCCCAACCATTTGATGCCACTCTGTATGATCACTACCAATCTTGGCGCA 591
 Db 165 PheMetGlyAlaPhePro----- 170
 QY 592 GAGGCGCTTGGGAGATCTCTGAAGATTCTCTTAGTGGACTACCTAGCTTGGACTTAGG 651
 Db 171 -----Gln 171
 QY 652 GTCCAGCTTAAGACTTCAGTCTATGACGAGGAGCTATGGAGGGCACTATGTCTCTGCA 711
 Db 172 TyrSerArgGluThrPheHisPheThrThrGluSerTyrGlyGlyHisTyrGlyProVal 191
 QY 712 TTCTTCAATCATTTTACGAGCAGAGATGAGAGAATTCGCCAAGGTAGTGTAAATGGTGT 771
 Db 192 PheAsnGluTyrIleGluGlnAsnAlaHisLeuGlnProGlyAla-----Lys 208
 QY 772 CAGCTTAATTTCAACTCTCTGGGAATATTAAACGGCATCATCAGCAGGGCATCCAGGCC 831
 Db 209 LysIleGlnLeuGlySerValMetIleGlyAsnGlyTrpTyrAspProIleIleGlnTyr 228
 QY 832 CCTTACTACCTGAATTCGGCTGTG-----AACAAATACCTACGGTATCAAGGTGTCAAC 885
 Db 229 GlnAlaTyrTyrAsnPheThrValTyrProGlyAsnThrTyrAspTyrLeuProPheAsn 248
 QY 886 GAGACCGCTTACAACTACATGAAGTTTGCACCAAAATGCCAAATGGTTGCCAGGATTG 945
 Db 249 LysSerIleSerSerLeuMetTyrAsnAsnLeuTyrGlyProGlyAsnCysLeuAspGln 268
 QY 946 ATTTCCACCTGCAACAGACAAACCCGACCGCATAGCTAGCTACGCCCTCTGCGCCGAA 1005
 Db 269 LeuTyrAspCys-----AlaAlaArgGlyIleAspGluIleCysSerThr 283
 QY 1006 GCCACCAACATGTCCAGGGACAATGTTGAGGGGCCATACCTAGCGCTTTGTGGTCTGGT 1065
 Db 284 AlaAspAspPheCysAlaAsnGluValGluAsnValTyrAspIleTyrSerGlyArgAsp 303
 QY 1066 GTGTATGATATTCGGCATCCATATGATGACCCGACTCCGCCAAGTTATTACACAAATTT 1125
 Db 304 GluTyrAspPheArgGluLeuThrProAspProPheProTyrGluPheTyrValAspTyr 323
 QY 1126 CTGCAAGAGACTCTGTCTATCGGCGTATCGGCTCAACACTCAACTCACTCACCAGTCCAAT 1185
 Db 324 LeuAsnLysAlaSerValGlnAlaAlaIleGlyAlaTyrIleAsnTyrThrGluSerAsn 343

QY 523 TTTCAGGAGTTCAGGCGCGGTACCCACCATTTGATGACCTCTGATCATCTACTACCAAT 582
 Db 182 -----GlulysValSerThrLys 188
 QY 583 CTTGCCGAGAGCGCTTGGGAGATCTTCAAGATTCCTTAGTGGACTACCTAGCTTG 642
 Db 189 LeuAlaGlyLysAlaTyrPheLeuGluLeuPheGluAlaPheProHisLeu 208
 QY 643 GACTCTAGGCTCAGTCTAAGACTTCAATCTTACGAGAGAGCTATGAGGCGACTAT 702
 Db 209 -----ArgSerAsnAspPheHisIleAlaGlyLysTyrAlaGlyHisTyr 224
 QY 703 GCTCTGCTCATCTTCAATCATTTTACGAGAGAT---GAGAGATTCGCCAAGCTAGT 759
 Db 225 IleProGlnIleAlaHisGluIleValValLysAsnProGluArg----- 239
 QY 760 GTTATGGTGTTCAGCTTATTTCACTCTCTGGGAATTTAAGCGCATCATCGACGAG 819
 Db 240 -----ThrPheAsnLeuThrSerValMetIleGlyAsnGlyIleThrAspPro 255
 QY 820 GCGATCCAGCGCCCTTACTACCTGAATTCGCTGTG---AACAAATACCTACGATCAAG 876
 Db 256 LeuIleGlnAlaAspTyrTyrGluProMetAlaCysGlyLysGlyTyrHisProVal 275
 QY 877 GCTGTCAACGAGACCGCTTACAACTACATGAGTTTGGCAACCAATGCAATGTTGC 936
 Db 276 LeuSerSerGluGluCysGluLysMetSerLysAlaAlaGlyArg----- 290
 QY 937 CAGGATTTGATTTCCACCTGCAACAGACAAACCGCACCGCATTTAGCTACGACGCTC 996
 Db 291 -----CysArgArgLeuAsnLysLeuCysTyrAlaSerLysSerSer 304
 QY 997 -----TGGCGCCAGCCCAACATGTGCGAGGCAATTTGAGGGGCCATCTACGCGC 1050
 Db 305 LeuProCysIleValAlaThrAlaTyrCysAspSerAlaLeuLeuGluProTyrIleAsn 324
 QY 1051 TTGCTGGTGGTGTATGATATTTGGCATCTCATATGATGACCGG----- 1098
 Db 325 ---ThrGlyLeuAsnValTyrAspIleArgGlyProCysGluAspAsnSerThrAspGly 343
 QY 1099 -----ACTCGGCCAGTTTATACAAATTTCTGCAAGAGACTCTGTGTCATGGAC 1149
 Db 344 MetCysTyrThrGlyLeuArgTyrValAspGlnTyrMetAsnPheProGluValGlnGlu 363
 QY 1150 GCTATCGGCTCAACATC---AACTACACCCAGCTCCCAATATACCTCTACTACGTTTC 1206
 Db 364 ThrLeuGlySerAspValHisAsnTyrSerGlyCysAspAsnAspValPheThrGlyPhe 383
 QY 1207 CAGCAACAGCGGACTTTGCTGGCCCACTTATCGAACACCTCGAGGAGATCTTGCT 1266
 Db 384 LeuPheThrGlyAspGlySerLysPro---PheGlnGlnTyrIleAlaGluLeuAsn 402
 QY 1267 CTCCCGTGGTGTCTCCCTCATCTATGGCGAGCCGATTAACATCTGCTGCTTCGCG 1326
 Db 403 HisAsnIleProValLeuIleTyrAlaGlyAspLysAspTyrIleCysAsnTrpLeuGly 422
 QY 1327 GGTGAGCGGCTTTCCCTCGTGGCACTTCTCCCAAGCCCGCCAGTTCGGAAGCGCAGG 1386
 Db 423 AsnHisAlaTrpSerAsnGluLeuGluTyrIleAsnLysArgArgTyrGlnArgMet 442
 QY 1387 TACAGCGCC-----CTGAAGTCAAGCGGTCTAGTATGGGAACTCGCGAGTATGGT 1440
 Db 443 LeuArgProTrpValSerLysGluThrGlyGluGluLeuGlyGlnValLysAsnTyrGly 462
 QY 1441 AATTTCCTCTCACTCGCTTATGAGCGAGCGCATGAAGTCCCATCTACACGCCCATC 1500
 Db 463 ProPheThrPheLeuArgIleTyrAspAlaGlyHisMetValProTyrAspGlnProGlu 482
 QY 1501 GCCTCCCTGCAATGTTTAAACCGGACTATCTTCGGT 1536
 Db 483 AlaSerLeuGluMetValAsnSerTrpIleSerGly 494

RESULT 7

ABR38819
 ID ABR38819 standard; Protein; 536 AA.
 XX
 AC ABR38819;
 DT 24-APR-2003 (first entry)
 XX
 DE A. niger carboxypeptidase Y polypeptide #1.
 XX
 KW protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.5.
 XX
 OS Aspergillus niger.
 XX
 PN WO200268623-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 22-FEB-2002; 2002WO-EP01984.
 XX
 PR 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0000075.
 PR 28-MAR-2001; 2001EP-0000078.
 PR 28-MAR-2001; 2001EP-0000080.
 PR 28-MAR-2001; 2001EP-0000087.
 PR 28-MAR-2001; 2001EP-0000088.
 PR 21-MAY-2001; 2001EP-0000156.
 PR 21-MAY-2001; 2001EP-0000159.
 PR 21-MAY-2001; 2001EP-0000160.
 PR 21-MAY-2001; 2001EP-0000162.
 PR 21-MAY-2001; 2001EP-0000165.
 PR 21-MAY-2001; 2001EP-0000166.
 PR 21-MAY-2001; 2001EP-0000168.
 PR 21-JUN-2001; 2001EP-0000240.
 PR 21-JUN-2001; 2001EP-0000242.
 PR 21-JUN-2001; 2001EP-0000244.
 PR 21-JUN-2001; 2001EP-0000246.
 PR 12-JUL-2001; 2001EP-0000280.
 PR 12-JUL-2001; 2001EP-0000285.
 PR 30-JUL-2001; 2001EP-0000323.
 PR 30-JUL-2001; 2001EP-0000327.
 PR 02-AUG-2001; 2001EP-0000341.
 PR 02-AUG-2001; 2001EP-0000342.
 PR 02-AUG-2001; 2001EP-0000343.
 PR 02-AUG-2001; 2001EP-0000344.
 PR 09-AUG-2001; 2001EP-0000357.
 PR 16-AUG-2001; 2001EP-0000374.
 PR 16-AUG-2001; 2001EP-0000377.
 PR 20-SEP-2001; 2001EP-0000478.
 PR 20-SEP-2001; 2001EP-0000483.
 PR 22-OCT-2001; 2001EP-0000552.
 PR 22-OCT-2001; 2001EP-0000553.
 PR 22-OCT-2001; 2001EP-0000554.
 PR 22-OCT-2001; 2001EP-0000556.
 PR 22-OCT-2001; 2001EP-0000557.
 PR 22-OCT-2001; 2001EP-0000558.
 PR 15-NOV-2001; 2001EP-0000464.
 PR 21-DEC-2001; 2001EP-00005117.
 XX
 PA (STAM) DSM NV.
 XX
 EI Edens L, Van Dijk AA, Krubasik P, Albertmann K, Stock A, Kimpel E;
 EI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 EI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 EI Stiebler J, Albarg R;
 XX

Db 254 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 273
QY 730 GAGCAAGAATGAGAGAAATGCCAAACGGTAGTGTAAATGGTGTTCACGTTAAATTCACACTCT 789
Db 274 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 284
QY 790 CTGGGAATATTAAACGGCATCATCAAGAGCGATCCAGGCCCTTACTACCTCGAATTC 849
Db 285 ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet 304
QY 850 GCTGTGAACAATACCTACGGTATCAAGGCTCTCAACGAGACCGCTCTACAACTACATGAAG 909
Db 305 AlaCysGlyAspGlyGlyTyrProAlaValLeuAspGluSerSerCysGlnSerMetAsp 324
QY 910 TTTGCCAACCAAAATGCCAAATGGTTCAGGATTTGATTTCCACTGCACACACACAAC 969
Db 325 -----AsnAlaLeuProArg-----CysGlnSerMetIleGluSerCysTyrSerSerGlu 349
QY 970 CGCACCAGATTAGCTAGCTACGCCCTCTGCGCGAAGCCACCAACATGTGCGAGGACAAT 1029
Db 342 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 356
QY 1030 GTTCAGGGGCCATACTACGCCCTTGTCTGGTCTGGTGTGTATGATATTCGGCATCCATAT 1089
Db 357 LeuLeuAlaProTyrGlnArg-----ThrGlyGlnAsnValTyrAspValArgGlyLysCys 375
QY 1090 GATGACCCG-----ACTCCGCGAAGTTATTACAAACAATTTCTGGCAAG 1134
Db 376 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 395
QY 1135 GACTCTGTCATGAGCGTATCGGCGTCAACATCAAC---TACACCCAGTCCCAATAATGAC 1191
Db 396 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 415
QY 1192 GTCTACTACGTTTCCAGCAACAGCGGACTTTGCTCTGGGCC-----AACTTCATCGAA 1245
Db 416 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 435
QY 1246 GACCTCGAGGAGATCTCTGTCTCCCGTCCGCGTCTCCCTCATCTAT---GGCGACGCC 1302
Db 436 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 450
QY 1303 GATTACATCTGCAACTGCTGCGCGGTGAGCGGTTCCCTCGCTCGCAACTACTCCCAA 1362
Db 451 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTyrProGly 470
QY 1363 GCGCCCGCAGTTCGAGCGGAGGTACACGCCCTCGAAAGTC-----AAC 1407
Db 471 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 490
QY 1408 GCGCTCGAGTATGGGAAACTCGCAGTAGTAGTTCTCTCTTCTACTCGCGTCTATGAG 1467
Db 491 GlyLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 510
QY 1468 GCAGGCCATCAATCCCTACTACAGGCCATCGCCTCCCTCGCAATGTTTAAACCGGACT 1527
Db 511 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 530
QY 1528 ATCTTCGGT---TGG 1539
Db 531 LeuGlyGlyGluTrp 535

RESULT 8
AAR96737
ID AAR96737 standard; Protein; 556 AA.
XX
XX AAR96737;
XX
DT 14-AUG-1996 (first entry)
DE
XX A. niger Bo-1 carboxypeptidase Y.
XX

Db 254 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 273
QY 730 GAGCAAGAATGAGAGAAATGCCAAACGGTAGTGTAAATGGTGTTCACGTTAAATTCACACTCT 789
Db 274 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 284
QY 790 CTGGGAATATTAAACGGCATCATCAAGAGCGATCCAGGCCCTTACTACCTCGAATTC 849
Db 285 ValIleuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet 304
QY 850 GCTGTGAACAATACCTACGGTATCAAGGCTCTCAACGAGACCGCTCTACAACTACATGAAG 909
Db 305 AlaCysGlyAspGlyGlyTyrProAlaValLeuAspGluSerCysGlnSerMetAsp 324
QY 910 TTTGCCAACCAAAATGCCAAATGGTTCAGGATTTGATTTCCACTGCACACACACAAC 969
Db 325 -----AsnAlaLeuProArg-----CysGlnSerMetIleGluSerCysTyrSerSerGlu 349
QY 970 CGCACCGCATTAGCTAGCTACGCCCTCTGCGCGAAGCCACCAACATGTGCGAGGACAAT 1029
Db 342 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 356
QY 1030 GTTCAGGGGCCATACTACGCCCTTGTGCTGTCGTGTGTATGATGATATTCGGCATCATAT 1089
Db 357 LeuLeuAlaProTyrGlnArg-----ThrGlyGlnAsnValTyrAspValArgGlyLysCys 375
QY 1090 GATGACCCG-----ACTCCGCGAAGTTATTACAAACAATTTCTGGCAAG 1134
Db 376 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 395
QY 1135 GACTCTGTCATGAGCGTATCGGCGTCAACATCAAC---TACACCCAGTCCCAATAATGAC 1191
Db 396 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 415
QY 1192 GTCTACTACGTTTCCAGCAACAGCGGACTTTGCTCTGGGCC-----AACTTCATCGAA 1245
Db 416 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 435
QY 1246 GACCTCGAGGAGATCTCTGCTCTCCCGTCGCGTGTCTCCCTCATCTAT---GGCGACGCC 1302
Db 436 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 450
QY 1303 GAITACATCTGCAACTGCTGCGCGGTGAGCGGTTCCCTCGCTCGCAACTACTCCCAA 1362
Db 451 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTyrProGly 470
QY 1363 GCGCCCGCAGTTCGAGCGGAGGTACACGCCCTCGAAAGTC-----AAC 1407
Db 471 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 490
QY 1408 GCGCTCGAGTATGGGAAACTCGCAGTAGTAGTTCTCTCTTCTACTCGCGTCTATGAG 1467
Db 491 GlyLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 510
QY 1468 GCAGGCCATCAATCCCTACTACAGGCCATCGCCTCCCTGCAATGTTTAAACCGGACT 1527
Db 511 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 530
QY 1528 ATCTTCGGT---TGG 1539
Db 531 LeuGlyGlyGluTrp 535

RESULT 8
AAR96737
ID AAR96737 standard; Protein; 556 AA.
XX
XX AAR96737;
XX
DT 14-AUG-1996 (first entry)
DE
XX A. niger Bo-1 carboxypeptidase Y.
XX

KW Carboxypeptidase Y; CPY; ascomycete; deuteromycete; host cell;
 KW protease deficiency.
 XX Aspergillus niger strain Bo-1.
 XX OS
 XX WO9609397-Al.
 XX PN
 XX 28-MAR-1996.
 XX PD
 XX 19-SEP-1995; 95WO-US11945.
 XX PF
 XX 20-SEP-1994; 94US-0309341.
 XX PR
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX PA
 XX Thompson SA, Yaver DS;
 XX PI
 XX WPI; 1996-108458/19.
 XX DR
 XX N-PSDB; AAT28283.
 XX PT
 XX Nucleic acid construct encoding a filamentous ascomycete or
 XX deuteromycete carboxypeptidase Y - useful to produce host cells
 XX modified to produce reduced amounts of carboxypeptidase
 XX PT
 XX Claim 4; Fig 1A-E; 46pp; English.
 XX PS
 XX Carboxypeptidase Y (AAR96737), a vacuolar protease, is the product of
 XX the CPY gene (AAT28283) of Aspergillus niger strain Bo-1. Creation
 XX of CPY-deficient Aspergillus strains, e.g. by cloning a selectable
 XX marker into the CPY gene, provides suitable host strains for prodn.
 XX of heterologous protein.
 XX CC
 XX SQ Sequence 556 AA;

Alignment Scores:
 Pred. No.: 7.22e-48 Length: 556
 Score: 630.00 Matches: 154
 Percent Similarity: 48.45% Conservative: 81
 Best Local Similarity: 31.75% Mismatches: 174
 Query Match: 20.81% Indels: 76
 DB: 17 Gaps: 17

US-09-712-338-1 (1-1668) x AAR96737 (1-556)

QY 133 GCAACAAATGTCACCATCCGGTACAGGAACCCGGGGCAGAGCGCTCTGCGAGCTACC 192
 DB 131 AlatyAspLeuArgValLysLysThrAspProGlySerLeuGlyIle-----Asp 147
 QY 193 CCGGTGTCAATCCTACTCTGGATATGTCACACCTCTCCC---GAGTCCCATACCTTC 249
 DB 148 ProGlyValLysGlnTyrThrGlyTyrLeuAspAspAsnGluAsnAspLysHisLeuPhe 167
 QY 250 TTCCTGGTCTTCGAAGCCAGACATAACCCGAACTGCACCTATCACATTTGGTGAAT 309
 DB 168 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 187
 QY 310 GGTGGCCCTGGAAGCGATCTTTGATCGGTCTCTCGAAGAGTTGGGCCCTTGGCATGTC 369
 DB 188 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile 207
 QY 370 AATTCGACTTTTGATGACTATCATACCTCCTGTCGAGAGGTTCCCAATTTACTA 429
 DB 208 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle 227
 QY 430 TTCCTGTCCAGCCATCGGAGTCGGCTTTTCATAGTATGATGATGGTTGATGGGTCCATT 489
 DB 228 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----243
 QY 490 AACCCGTGTAACCTGGGGTCTCGAAAAATTCGAGCTTTGCGAGGAGTTTCAGGGCCGTACCCA 549
 DB 243 -----243
 QY 550 ACCATTGATGCCACTCTGATCGATACCTACCANTCTTGGCGAGAGCGCTTGGGAGATC 609

DB 244 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 257
 QY 610 CTGCAAGGATTCCTTAGTGGACTACTAGCTGGACTCTAGGGTGCAGTCTAGGACTTC 669
 DB 258 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 273
 QY 670 AGTCTATGACGAGAGCTATGAGGCACTATGGCTCTGCATCTTCAATCATTTTAC 729
 DB 274 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 293
 QY 730 GAGCAGATGAGAGAAATCCCAACGGTAGTGTAAATGGTGTTCAGCTTAATTTCAACTCT 789
 DB 294 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 304
 QY 790 CTGGGAATTTAAGGCATCATCGACGAGGCGATCCAGGCCCTTACTACCTCGAATTC 849
 DB 305 ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet 324
 QY 850 GCTGTGAACAATACCTACGGTATCAAGGCTGTCACGAGACCGCTCTACAACTACATGAAG 909
 DB 325 AlaCysGlyAspGlyGlyTyrProAlaValLeuAspGluSerSerCysGlnSerMetAsp 344
 QY 910 TTTCGCCAACCAATGCGCAATGTTGCCAGGATTTGATTTCCACCTGCCAAACAGACAAAC 969
 DB 345 -----AsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 361
 QY 970 CGCACCGCATTTAGCTAGCTAGCCCTCTGCGCGAGACCCACCAACATGTCAGGGACAAT 1029
 DB 362 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 376
 QY 1030 GTTGAGGGGCGCATACTACGGCTTGTGGTGGTGTGTATGATATTCGGCATCCATAT 1089
 DB 377 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 395
 QY 1090 GATGACCCG-----ACTCCGCCAAGTTATTACAACTTTCTGGCAAG 1134
 DB 396 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 415
 QY 1135 GACTGTGTATGAGACGTATCGGGTCAACATCAAC---TACACCCAGTCCATAATGAC 1191
 DB 416 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 435
 QY 1192 GTCTACTACGCTTCCAGCAACAGCGAGCTTTGTCTGGCCC-----AACTTCATCGAA 1245
 DB 436 IleAsnArgAsnPheLeuPheHisGlyAspIlePheMetLysProTyrHisArgLeuValPro 455
 QY 1246 GACCTCGAGGAGATCTTGTCTCTCCCGTGGGTGTCTCCCTCACTAT---GGCGAGCC 1302
 DB 456 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 470
 QY 1303 GATTACATCTGCAACTGTTGCGGGGTGCGGCGCTTCCCTCGTGGAACTACTCCCA 1362
 DB 471 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 490
 QY 1363 GCGCCCGAGTTCGGAAGCGAGGTTACAGCGCCCTGAAAGTC-----AAC 1407
 DB 491 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 510
 QY 1408 GCGCTCGAGTATGGGAAACTCGCGAGTATGTTGTAATTTCTCTTCACTCGCGTCTATGAG 1467
 DB 511 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 530
 QY 1468 GCAGGCCATGAGTCCCATCTACTACAGCCCATCCCTCCCTGCAATGTTTAAACCGGACT 1527
 DB 531 GlyGlyHisMetValProMetAspGlnProGluSerLeuGluPhePheAsnArgTrp 550
 QY 1528 ATCTTCGCT---TGG 1539
 DB 551 LeuGlyGlyGluTrp 555
 RESULT 9
 AAR96738

[illegible]

QY 1408 GCGTCGAGTATGGGAACTCCGAGTATGTAATTTCTCTACTCGGTCATAGAG 1467
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsn2PheThrPheMetArgLeuTyrGly 531
 QY 1468 GCAGGCATCAAGTCCCTACTACACGACCATCGCTCCCTGCAATTGTTTAAACGGACT 1527
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTyr 551
 QY 1528 ATCTTCGGT---TGG 1539
 Db : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 552 LeuGlyGlyGluTyr 556

RESULT 10
 AAR48059
 ID AAR48059 standard; Protein; 491 AA.
 XX
 AC AAR48059;
 XX
 DT 25-MAR-2003 (updated)
 DT 20-JUL-1994 (first entry)
 XX
 DE Sequence of protease C encoded by gene K1.PRC1.
 XX
 KW Protease; yeast; proteolysis.
 XX
 OS Kluyveromyces lactis.
 XX
 PN WO9400579-A1.
 XX
 PD 06-JAN-1994.
 XX
 XX 23-JUN-1993; 93WO-FR00623.
 XX
 XX 25-JUN-1992; 92PR-0007785.
 XX
 XX (RHON) RHONE POULENC RORER SA.
 PA
 XX Flier R, Fournier A, Yeh P;
 PI
 XX WPI; 1994-026215/03.
 DR
 DR N-PSDB; AAQ55347.
 DR
 XX New Kluyveromyces yeast with modified protease gene - esp. used
 PT for high yield prodn. of recombinant protein, also DNA encoding
 PT yeast protease and derived peptide(s)
 XX
 XX Disclosure; Page 28-31; 49pp; English.
 PS
 XX The protease gene is to be modified in order to render it (partially)
 CC incapable of producing the natural protein; or result in a non-
 CC functional protease or in a protease with modified proteolytic
 CC activity. The modifications can be introduced in vitro or in situ by
 CC standard genetic engineering techniques or by exposure to mutagenic
 CC agents.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 491 AA;

Alignment Scores:
 Pred. No.: 2,44e-42 Length: 491
 Score: 568.50 Matches: 151
 Percent Similarity: 44.65% Conservative: 66
 Best Local Similarity: 31.07% Mismatches: 184
 Query Match: 18.78% Indels: 85
 DB: 15 Gaps: 14

US-09-712-338-1 (1-1668) x AAR48059 (1-491)

QY 133 GCAACATGTCACATCCGCTACAGGAACCCGGGGCAGAGGGGCTCTCGAGACTACC 192
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 62 AlaTyrSerLeuArgLysProLeuAspProLysSerLeuGlyVal----- 77
 QY 193 CCGGGTGTCAATCTACTCTGATATGTGACACACCTCTCCGAGTCCCATACCTTCTTC 252

Db 78 AspThrValLysGlnTyrSerGlyTyrLeuAspTyrGlnAspSerLysHisPhePheTyr 97
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 253 TGGTCTTTCGAGCCAGACATAACCCAGAACTGCCTATACCATTTGCTGTTGAATGGT 312
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 98 TrpPhePheGluSerArgAsnAspProGluAsnAspProValIleLeuTyrLeuAsnGly 117
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 313 GGCCTGGAAGGATCTCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTGCATGTCAT 372
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 118 GlyProGlyCysSerSerPheValGlyLeuPheGluLeuGlyProSerSerIleGly 137
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 373 TCGACTTTTGATGACTACATCAACCTCTCTGGAACGAGGTCTCCAAATTTACTATT 432
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 138 AlaAspLeuLysProLysTyrAsnProTyrSerIlePheAsnSerAsnAlaSerValIlePhe 157
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 433 CTGTCCCGACCATTTGGAGTCGGCTTTTCATATAGTATACGGTTGATGGTCCATTAAC 492
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 158 LeuAspGlnProValGlyValGlyPheSerTyrGlyAsp----- 170
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 493 CCTGTAACTGGGTCTGCGAAATTCGAGCTTTGAGGAGTTGAGGCCGGTACCCCAACC 552
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 170 ----- 170
 QY 553 ATTGATGCCACTCTGATCGATACCTACCAATCTTCCGCGAGAGCCGCTTGGAGATCCTG 612
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 171 -----SerLysValSerThrThrAspAspAlaLysAspValTyrIlePheLeu 187
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 613 CAAGGATTCCTTAGTGGACTACCTAGCTTGGACTCTAGGTGAGTCTAAGAGCTTCAGT 672
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 188 AspLeuPhePheGluArgPheProHisLeu-----ArgAsnAspPheHis 203
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 673 CTATGACGAGGAGCTATGAGGGGACTATGCTCTGCTCATCTTCAATCATTTTACGAG 732
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 204 IleSerGlyGluSerTyrAlaGlyHisTyrLeuProLysIleAlaHis----- 219
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 733 CAGAATGAGAGAATTGCCAACGGTGTGTTAATGTTTCCAGCTTAATTTCAACTCTCTG 792
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 220 -----GluIleAlaValValHisAlaGluAspSerSerPheAsnLeuSerVal 236
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 793 GGAATTATTAAACGGCATCATCGACGAGCGGATCGAGCCCTTACTACCTGTAATTCGCT 852
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 237 LeuIleGlyAsnGlyPheThrAspProLeuThrGlnTyrGlnTyrGluProMetAla 256
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 853 GTGAACATACCTACGGTATCAAGGTGTCAAGAGACCGCTCTACACTACATGAATTT 912
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 257 CysGlyGluGly---GlyTyrProAlaValLeuGlu----- 267
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 913 GCCAACCAATGCCAATGCTTCCAGGAT----- 942
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 268 -----ProGluAspCysLeuAspMetAsnArgAsnLeuProLeuCysLeuSer 283
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 943 TTGATTTCCACCTGCAACAGACAAACCCGACCGCATTAGCTAGTACGCTCTGGGCC 1002
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 284 LeuValAspArgCysTyrLysSerHis-----SerValPheSerCysVal 298
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 1003 GAAGCCACCAACATGTCAGGGGCAATGTTGAGGGCCATACGCTTTCGCTGTCGT 1062
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 299 LeuAlaAspArgTyrCysGluGlnIleThrGly---ValTyrGluLysSerGlyArg 317
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 1063 GGTGTGTATGATATTCGG-----CATCATATGATGACCG-----ACT 1101
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 318 AsnProTyrAspIleArgSerLysCysGluAlaGluAspAspSerGlyAlaCysTyrGln 337
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 1102 CCGCAAGTATTATACAAATTTCTGCAAGACCTCTGTCATGAGCGTATCGGCGTC 1161
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 338 GluGluIleTyrIleSerAspTyrLeuAsnGlnGluValGlnArgAlaLeuGlyThr 357
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 1162 AACATCAACTACACCCAGTCC---ATAATGACCTCTACTACGCTTTCAGCAACACAGGC 1218
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 358 AspValSerSerPheGlnGlyCysSerSerAspValGlyIleGlyPheAlaPheThrGly 377
 QY ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 1219 GACTTTGTGCGCCAACTTCATCATGAAGACTCTCGAGGAGATCTCTCCCGTTCGT 1278
 Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 1219 GACTTTGTGCGCCAACTTCATCATGAAGACTCTCGAGGAGATCTCTCCCGTTCGT 1278

PR	26-FEB-2001;	2001EP-0200708
PR	26-FEB-2001;	2001EP-0200708
PR	26-FEB-2001;	2001EP-0200719
PR	28-MAR-2001;	2001EP-0000075
PR	28-MAR-2001;	2001EP-0000078
PR	28-MAR-2001;	2001EP-0000080
PR	28-MAR-2001;	2001EP-0000087
PR	28-MAR-2001;	2001EP-0000088
PR	21-MAY-2001;	2001EP-0000136
PR	21-MAY-2001;	2001EP-0000159
PR	21-MAY-2001;	2001EP-0000160
PR	21-MAY-2001;	2001EP-0000162
PR	21-MAY-2001;	2001EP-0000165
PR	21-MAY-2001;	2001EP-0000166
PR	21-MAY-2001;	2001EP-0000168
PR	21-JUN-2001;	2001EP-0000240
PR	21-JUN-2001;	2001EP-0000242
PR	21-JUN-2001;	2001EP-0000244
PR	21-JUN-2001;	2001EP-0000246
PR	12-JUL-2001;	2001EP-0000280
PR	12-JUL-2001;	2001EP-0000285
PR	30-JUL-2001;	2001EP-0000323
PR	30-JUL-2001;	2001EP-0000327
PR	02-AUG-2001;	2001EP-0000341
PR	02-AUG-2001;	2001EP-0000342
PR	02-AUG-2001;	2001EP-0000343
PR	02-AUG-2001;	2001EP-0000344
PR	09-AUG-2001;	2001EP-0000357
PR	16-AUG-2001;	2001EP-0000374
PR	16-AUG-2001;	2001EP-0000377
PR	20-SEP-2001;	2001EP-0000478
PR	20-SEP-2001;	2001EP-0000483
PR	22-OCT-2001;	2001EP-0000552
PR	22-OCT-2001;	2001EP-0000553
PR	22-OCT-2001;	2001EP-0000554
PR	22-OCT-2001;	2001EP-0000556
PR	22-OCT-2001;	2001EP-0000557
PR	22-OCT-2001;	2001EP-0000558
PR	15-NOV-2001;	2001EP-0004464
PR	21-DEC-2001;	2001EP-0005117

(STAM) DSM NV.

Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E
Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
Stiebler J, Albang R

WPT: 2002-723203/78.

WPI; 2002-723203/78.
P-PSDB: ABZ78212. ABZ78269.

Novel isolated protease polypeptide useful in laboratory, clinical, pharmaceutical, chemical, diagnostic, personal care and industrial applications -

Claim 13: page 335-337: English: English.

The invention relates to a novel isolated protease polypeptide. A polypeptide or polynucleotide of the invention is useful for diagnosing a fungal infection such as aspergillosis, or as a query sequence to perform a search against public databases. A polypeptide of the invention is useful in a selected number of industrial or pharmaceutical processes, in laboratory or clinical processes, in food industry (baking, brewing, cheese manufacturing, meat tenderising), in tanning industry and in the manufacture of biological detergents. A polypeptide may also be useful for improving protein solubility, extraction yields, viscosity or taste, texture, nutritional value, minimising of antigenicity or anti-nutritional factors, colour or functionality as well as processing aspects like filterability of the proteinaceous raw material. The sequences shown in ABR38313-ABR38369 represent the *A. niger* proteases of the invention.

Sequence 481 AA;

OS
XY

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0126234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148119.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151103.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0158369.

XX AC AAG30064;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 35876.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

Db 51 eCysAspAlaArgSer-----ProGlnPheThrGlyTrpLeuAspIleGlyProLys-- 68
QY 240 CCATACCTTCCTCTGCTTCTTCGAAGCCAGACATAACCCAGAACTGCACCTATCACATT 299
Db 69 -HisLeuPhePheTrpIleGluSerGlnAspProPheHisAspProLeuThrLe 88
QY 300 GTGGTTGAAGTGGCCCTCGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCC 359
Db 88 uTrpMetThrGlyGlyProGlyAspSerSerMetIleGlyLeuPheGluGluValGlyPr 108
QY 360 TTGCATGTCNAAT-----TCGACTTTTGATGACTACATCAACCTCTACTCGTGAAGA 413
Db 108 oCysArgIleAsnGluPheGlyAsnGlyThrAspHis---AsnProTrpAlaTrpThrLy 127
QY 414 GGTCTCCAATTACTATTCCNGTCCCGAGCTCCAGCCATTGGAGTCGGCTTTCATATAGTGATAC 473
Db 127 sAsnSerSerLeuLeuPheValAspGlnProValAspValGlyPheSerIyrIleAsp-- 146
QY 474 GGTGATGGTCCATTAAACCTGTAACTGGGGTCGTCGAAAAATTCGAGGAGT 533
Db 147 -----GluGlyTy 149
QY 534 TCAGGGCGGTACCCAAACCATTTGATGCCACTCTGATCGATCTACTACCAATCTTGGCGAGA 593
Db 149 rGlu-----LeuProHisAspSerArgGluAlaAlaVa 160
QY 594 GGCCGCTTGGGAGATCTCGAAGGATTCCTTAGTGGACTACCTAGCTGGACTCTAGGGT 653
Db 160 lAspMetHisArgPheLeuArgLeuPheIleSer-----GluIlePheProHisLy 177
QY 654 GCAGTCTAAGGACTTCACTCTATGACGGAGAGCTATGGAGGCACTATGGTCTCTGCATT 713
Db 177 sGlnPheLeuProValHisLeuSerGlyGluSerIyrAlaGlyArgIyrIleProTyrLe 197
QY 714 CTTCAATCATTTTACGNGCAGATGAG----- 741
Db 197 uAlaThrGlnIleLeuGluGlnAsnGluLeuTyrLysAspSerProArgIleProLeuLy 217
QY 742 -----AGATTGCCAACGCTAGTGTAAATGGTCTTACAGCTTAATTTCAACTCTCTGGG 794
Db 217 sSerCysLeuValGlyAsnGlyPheMetSerProLysAspAlaThrPheGly-TyrTrpG 237
QY 795 AATATTAAACGGCATCATCGAGCGGATCCAGGCCCTTACTACCGCTTACTACATGAAAGTTGC 854
Db 237 lThrLeuCysThrThrAsnSer-GlyValProSerProIle-Phe----- 251
QY 855 GAACAAATACCTACGGTATCAAGGTGTCAACGAGACCGTCTACAACTACATGAAAGTTGC 914
Db 252 -----AsnGluThrArgCysAspIleMet-----Al 260
QY 915 CAACCAATGCCAAATGTTCCAGGATTTGATTTCCACCTGCAACACAGACAAACCCGCAC 974
Db 260 aAlaAsnMetProHis---CysMetAspLeuTyrAspIleCysIleGlnHis----- 276
QY 975 CGCATTAGCTACGCGCTCTCGCGGAGCCACCAACATGTGCGAGGGACAATGTGA 1034
Db 277 -----SerAspProAlaIleCysHisAlaAlaGlnSerValCysIyrAspSerValVa 294
QY 1035 GGGGCCATCTACGCTTGTGCTGGTGTGTGTATGATATTCGGCATCCATATGATGA 1094
Db 294 lGlyLeuMetAlaLysLeuLeuLeuArg-----MetThrTh 306
QY 1095 CCCGACTCCGCCA-----AGTTATTACAACA 1121
Db 306 rValThrAlaProCysGluIleAspGluMetCysTyrIleGluAlaAlaLeuIleGluAr 326
QY 1122 ATTTCTGGCAAGGACTCTGTCTATGCGAGCTATCGGCTCAACTCAACTACACCCAGTC 1181
Db 326 gTyrLeuAsnSerProSerValTrpGluAlaLeuSerProGlnGlnValThrGluTy 346
QY 1182 CAAT-----AATGACGTCTACTAGCTTTCCAGAAACAGCGGACTTTGTGCGCC 1232
Db 346 rLysPheValAlaThrSerValIleAspAlaPheAlaGlnSerAlaAsp----- 362

QY 1233 CAACTTCATCGAACACCTCGAGGAGATC-----CTTGCTCTCCCGCTGCGTGTCTC 1283
Db 363 -GlyMetValSerSerSerLysGlnIleAlaPheLeuLeuAlaAsnAsnValAspPheLe 382
QY 1284 CCTCATCTATGGCGACGCCGATTACATCTGCAACTGGTTCCGGCTCAGGCCGTTTCCT 1343
Db 382 uAlaTyrGlnGlyAsnLeuAspLeuAlaCysAsnThrAlaGlyAsnLeuArgTrpAlaAs 402
QY 1344 CGTTGCCAACTACTCCCAAGCCGCCAGTTCGGAAGCGCAGGTACAGCCCTCGAAAGT 1403
Db 402 nSerLeuSerTrpLysGlyGlnThrGluPheThrAlaLysProLeuLeuProTrpGluIl 422
QY 1404 CAACGGGCTCGAGTATGGGAA---ACTCGCGAGTATGGTAAATTTCTCTTCTCATCTCGGT 1460
Db 422 eGln---ValSerValGlyGluGlyThrAspGluThrSerArgPheAlaPheValThrVa 441
QY 1461 CTATGAGGCGACGCCCAT 1476
Db 441 lAspAsnAlaGlyHis 446

Search completed: November 21, 2003, 17:40:00
Job time : 86 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:55:56 ; Search time 39 seconds
(without alignments)
3672.287 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 2979

Sequence: 1 MRGYEFLSVLPVAASWALP.....HTQSSVPLPTATSMSSVGMA 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertibrate:*
- 14: sp.unclassified:*
- 15: sp.rvirus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2976	99.9	555	3 Q96VZ9	Q96VZ9 aspergillus
2	749	25.1	460	3 Q12569	Q12569 absidia zyc
3	601	20.2	552	3 Q96VC4	Q96VC4 emericella
4	596.5	20.0	541	3 Q94152	Q94152 pichia angu
5	595.5	20.0	537	3 Q14414	Q14414 pichia angu
6	479	16.1	507	10 Q8VWQ0	Q8VWQ0 gossypium h
7	473.5	15.9	501	10 Q9XH61	Q9XH61 matricaria
8	473	15.9	429	10 Q8GVT1	Q8GVT1 oryza sativ
9	466.5	15.7	505	10 Q9FFB0	Q9FFB0 arabidopsis
10	464.5	15.6	508	10 Q8L6A7	Q8L6A7 theobroma c
11	439	14.7	510	10 Q9LXH4	Q9LXH4 arabidopsis
12	438	14.7	510	10 Q93ZC3	Q93ZC3 arabidopsis
13	417.5	14.0	2105	5 Q17679	Q17679 caenorhabdi
14	411.5	13.8	499	10 Q9FMX9	Q9FMX9 arabidopsis
15	385	12.9	490	10 Q9FPY7	Q9FPY7 oryza sativ
16	383	12.9	482	10 Q9SV04	Q9SV04 arabidopsis

17	381	12.8	524	10 Q8GTK2	Q8GTK2 oryza sativ
18	364	12.2	480	10 Q9LEY1	Q9LEY1 arabidopsis
19	361	12.1	482	10 Q9FRJ0	Q9FRJ0 oryza sativ
20	360	12.1	437	10 Q9FWG1	Q9FWG1 oryza sativ
21	359	12.1	487	10 Q9SV03	Q9SV03 arabidopsis
22	357.5	12.0	479	10 Q94907	Q94907 arabidopsis
23	355.5	11.9	360	10 Q9M450	Q9M450 cicor arlet
24	353	11.8	470	10 Q22732	Q22732 arabidopsis
25	351	11.8	669	10 Q8LPY6	Q8LPY6 oryza sativ
26	349.5	11.7	474	11 Q9D2D1	Q9D2D1 mus musculu
27	348	11.7	452	10 Q9ZQ00	Q9ZQ00 arabidopsis
28	344	11.5	465	10 Q04084	Q04084 arabidopsis
29	344	11.5	494	10 Q9FEU4	Q9FEU4 pisum sativ
30	344	11.5	501	10 Q9SV02	Q9SV02 arabidopsis
31	344	11.5	512	5 Q76725	Q76725 caenorhabdi
32	342.5	11.5	471	10 Q8L9Y0	Q8L9Y0 arabidopsis
33	342.5	11.5	473	10 Q9SGA9	Q9SGA9 arabidopsis
34	337.5	11.3	425	10 Q65568	Q65568 arabidopsis
35	336	11.3	465	10 Q9M099	Q9M099 arabidopsis
36	335.5	11.3	497	10 Q8L7B2	Q8L7B2 arabidopsis
37	333.5	11.2	510	3 Q60123	Q60123 schizosacch
38	333	11.2	487	10 Q9ZUG3	Q9ZUG3 arabidopsis
39	333	11.2	548	4 Q9BR08	Q9BR08 homo sapien
40	331	11.1	465	10 Q94K84	Q94K84 arabidopsis
41	330.5	11.1	472	10 Q9LSM9	Q9LSM9 arabidopsis
42	330.5	11.1	474	10 Q82228	Q82228 arabidopsis
43	330	11.1	445	5 Q814B3	Q814B3 caenorhabdi
44	330	11.1	1203	5 Q45916	Q45916 caenorhabdi
45	328	11.0	504	10 Q9LSV8	Q9LSV8 arabidopsis

ALIGNMENTS

RESULT 1

Q96VZ9 ID Q96VZ9 PRELIMINARY; PRT; 555 AA.
AC Q96VZ9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Carboxypeptidase S1 precursor (EC 3.4.16.6).
GN CPI.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK3;
RA van den Broek P.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394242; AAK77166.1;
DR InterPro; IPR001563; Serine carbpept.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PRODEM; PD001189; Serine carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase; Hydrolase; Signal.
FT SIGNAL 1
FT POTENTIAL 18
SQ SEQUENCE 555 AA; 61168 MW; 456B63B0CEB55222 CRC64;

Query Match 99.9%; Score 2976; DB 3; Length 555;
Best Local Similarity 99.8%; Pred. No. 7e-201;
Matches 554; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGYEFLSVLPVAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRYKEPGAEGV 60
Db 1 MRGYEFLSVLPVAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRYKEPGAEGV 60
QY 61 CETTGVKSYSGYVDTSPESHFTFFWFFEARHNPETAPITLWNGGPGSDSLGLFEELGP 120
Db 61 CETTGVKSYSGYVDTSPESHFTFFWFFEARHNPETAPITLWNGGPGSDSLGLFEELGP 120

QY	121	CHVNSTFDYINPHSNWEVSNLLFLSQPLGFGFSYSDTVDGSINPVTGVVENSSEFAGVQG	180
Db	121	CHVNSTFDYINPHSNWEVSNLLFLSQPLGFGFSYNDTVDGSINPVTGVVENSSEFAGVQG	180
QY	181	RYPTIDATLIDTNNLAAEAAWEILQGLFSLGSLDSRVQSKDFSLMTESYGHGYGPAFN	240
Db	181	RYPTIDATLIDTNNLAAEAAWEILQGLFSLGSLDSRVQSKDFSLMTESYGHGYGPAFN	240
QY	241	HYEQNERIANGSVNGVOLNFSNLGINIIDIIEAIQAPYYPEFAVNNTYGIKAVNETVN	300
Db	241	HYEQNERIANGSVNGVOLNFSNLGINIIDIIEAIQAPYYPEFAVNNTYGIKAVNETVN	300
QY	301	YMKFANQWPNQGODLISLTCQNTNRALTADYALCAEATNMCRDNVEGPPYAFAGRGVYDIR	360
Db	301	YMKFANQWPNQGODLISLTCQNTNRALTADYALCAEATNMCRDNVEGPPYAFAGRGVYDIR	360
QY	361	HPYDDPTPPSYNKNFLAKDSVMDAIGVNNINTQSNNDVYYAFQOTGDFVWPNFIEDLEI	420
Db	361	HPYDDPTPPSYNKNFLAKDSVMDAIGVNNINTQSNNDVYYAFQOTGDFVWPNFIEDLEI	420
QY	421	LALPVRVSLIYGDAVYICNWFGGQAVSLAANYSAQAQFRSAGYTPLKVNGVEGETREVG	480
Db	421	LALPVRVSLIYGDAVYICNWFGGQAVSLAANYSAQAQFRSAGYTPLKVNGVEGETREVG	480
QY	481	NFSFTRVVEAGHEVPPYQPIASLQLFNRTIFGWDIAEGQKKIWPYSYKTNGTATATHTQSS	540
Db	481	NFSFTRVVEAGHEVPPYQPIASLQLFNRTIFGWDIAEGQKKIWPYSYKTNGTATATHTQSS	540
QY	541	VLPATATSMSSVGMA	555
Db	541	VLPATATSMSSVGMA	555

```

RESULT 2
Q12569 Q12569 PRELIMINARY; PRT; 460 AA.
ID AC Q12569;
AC Q12569;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE DE Prepro-carboxypeptidase Z (EC 3.4.16.5).
GN SPCZ.
OS Absidia zychae.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Absidia
OX NCBI_TaxID=36079;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NRIC 1199;
RC MEDLINE=95308538; PubMed=7788719;
RA Lee B.R., Takeuchi M., Kobayashi Y.;
RT "Molecular cloning and sequence analysis of the scpZ gene encoding the
RL Curr. Genet. 27:159-165(1995).
DR EMBL; D16519; BAA03966.1; -.
DR HSSP; P00729; 1YSC.
DR MEROPS; S10.010; -.
DR Interpro; IPR001563; Serine_carbpept.
DR Interpro; IPR000379; Ser_estrs site.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CR00XYPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
KW Carboxypeptidase; Hydrolase.
FT CHAIN 52 460
SQ SEQUENCE 460 AA; 50825 MW; CE3A410764548BB6 CRC64;

Query Match 25.1%; Score 749; DB 3; Length 460;
Best Local Similarity 35.7%; Pred. No. 1.9e-44;
Matches 179; Conservative 72; Mismatches 188; Indels 62; Gaps 15;

QY 12 IVAASWALPGSTPASVGRRQLPKNPCTGVKTLTTANNVTIRYKEPGARGVCETTPGVKSYS 71

```

```

18  LVCAPVTVOAHPMWHLVRQ-DGNDT-----SSGNTTQLKYTSP---KLCD--PDVKQYS 66
72  GYDTPESHTFFWFFPEARHNPETAPITLWLNGPGSDSLIGLFEELGPHCHVNSTFDDYI 131
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
67  GYLDAANDERHYFFWFESKNDKPNKDPLTITLWNGPGCCSLIGLWBELGPCQNGS----A 122
132 NPHSWNEVENLLFLSLOPLGVGFSYSTVDGSINPVTGVVENSFAGVQGRYPTIDATLID 191
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
123 NPHSWHNSNMULFFDQPDGVGFSY-----GKQ-----TVS 152
192 TTNLAAEAAWEILQGLFSLDSRVQSKDFSLMTESYGGHYGPAFFNFHYEQNERIAN 251
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
153 TTDEAERAWTFLQAYEYTFP---QYSKLDVHYFGESYGGHYIPGFASHVVDMMKKVQS 208
252 GSYNGVOLFNLSLIGTINGIIDEALQAPYPPEPAVNNTYGIKAVNEIVNYMKFANQMPNG 311
209 GEEKGVVPUKSTGVNGVGFIDAVIQYKSPKMTCHSTYPAVLSEE---ECDMMQQYEND 265
312 CQDLSTCKQTNRTALADYALCAEATNMCNDRNVEGVYAFAGRGVYDIRHYPDDTPPSY 371
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
266 CKFAAQOCASD-----EBSDCVNAQCQ-GEIEG-IYACSGYSFYDIRQQGDD-TPHPF 317
372 YNKFLAKSDYMDAIGNVINTYQSNNDVYVAFQOTGDFVWENFTLEDEILALPVRVSLIY 431
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
318 VDE--LNKASVIKEVGARGHFSGSDSVGTAFAGTGDGA-RSYTPAVEKLLKEGIPVLIVY 375
432 GDADYICNWFQGAVSLAANYSAAQFRSAGYTPFLKVGNGVEYGETREYGNFSTRVVEAG 491
376 GDADVICNTIGNLADVADSLKWGDSDAFSKTLEAWKADGKEVQCFRSADKLITFVRVVEAG 435
492 HEVPPYQPIASLQLENRTIFG 512
436 HEVPMYQPEAALSMTOTWISG 456

```

[illegible]

QY 64 FWFPEARHNPETAPITLWNGSGDSLSLIGLFEELGPGCHVNSTFDDYINPHSWNEVSNLL 143
Db 148 YWFEARNNDPVNDPVLWNGPGGSSLTGLMFLGSGASIGPDLKPINNPYSWNSNATVI 207
QY 144 FLSOPLGVGFYSVDTVDGINSNPVTVGVNSFAGVQGRYPIDATLIDTTLAAEAAWEI 203
Db 208 FLDQPVNVGFYSYS-----SKSVNTVAAGKDVYAF 237
QY 204 LOGELSGLPDSLRVQSKDFSLWTSYGSHYGPAPFNHFYQNERIANGSVNGVQLNENS 263
Db 238 LELFYQQPHL-----LKNDFHAGESYGHYIPVAFSAILTHADR-----SFNLT 284
QY 264 LGIINGIIDEATQAPYEFVANNYGIK-AVNETVYNYMKFANQMPGQDILSTCKQT 322
Db 285 VLIINGLTDPLNQYFYSERMASCTDGGVEPTLDESECEGM--LETLP-CLSLIESCYSS 341
QY 323 NRTALADVALCAEATMCRDNVEGYPYAFAGGVYDIRHP-----YDDPTPPSYNKFEL 376
Db 342 Q-----SVFSCVPASIYCNNAQLGPFQK-TGRNVYDVRKVCBGTLCYXD---MEYIDQYL 392
QY 377 AKDSYMDALGVNIN-YTQSNNDVYAFQQTGDEVAP---NFTEDLEEILALPVRVSLIY- 431
Db 393 NODFVKVGAEDVTYESCFDVRNRLFPAGDWMKPYHKNVNLLEQ--GLPV---LIYA 447
QY 432 GDADYICNWFQGGQAVSLAANYSOAAQFRSAG-YTELKVNVEYGETREYNFSFTRVYEA 490
Db 448 GDKDFICNWLGNQANSNLPWSGHDEFESAELYNLTLDGTGVKGVNAGKFTFARMFDG 507
QY 491 GHEVYPIQPIASLQFNRTIFGWDTAEGQK 521
Db 508 GHEVYDQPESSLAMVNRWIAG-DYSLGTGK 537

RESULT 6
Q8VWQO PRELIMINARY; PRT; 507 AA.
AC Q8VWQO;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative serine carboxypeptidase.
GN SCP.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=3635;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Xu-142; TISSUE=Fiber;
RC Li C.-H., Zhu Y.-Q., Meng Y.-L., Wang J.-W., Xu K.-X., Zhang T.-Z.,
RA Chen X.-Y.;
RT "Isolation of Genes Preferentially Expressed in Cotton Fiber Cells by
RT CDNA Array and RT-PCR."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072822; AAL67992.1; -
DR MEROPS; S10.009; -
DR InterPro; IPR001563; Serine_carbpept.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PROSITE; PS001189; Serine_carbpept; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 507 AA; 56594 MW; 6A0D283D9C5E7713 CRC64;
Query Match 16.1%; Score 479; DB 10; Length 507;
Best Local Similarity 27.4%; Pred. No. 2.le-25;
Matches 149; Conservative 90; Mismatches 205; Indels 100; Gaps 21;
QY 13 VAASWALPGSTPASVGRRLPKN---PTGVKLTITANNVTI-----RYKEPGAEG 59

Db 27 ITADLRITDSFPSPVHAKGLIRELNLPKGVNVVDEHRVSLPEGPKLVEKRFPTLE- 85
QY 60 VCETTPGVKSY-----SGY--VDTSPESTHFFWFPEARHNPETAPITLWNGPGSDSL 111
Db 86 ----VPGVSPEDLGHAGYKFLPNSHDARMEYFFFSR-NSKXDPVIVLITGGPGCSSE 140
QY 112 IGLFEELGPGCHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFYSVDTVDGINSNPVTVGV 171
Db 141 LALFVENGPFPIADNMSLVNWEYGMWKASNLLYVDQPIGTGFSYS----- 185
QY 172 NSSFAGVQGRYPITDITLIDTTLAAEAAWEILOGELSGLPDSLRVQSKDFSLWTSY 231
Db 186 --SDQDRIHNEDEVSNLDY-----FLQAFPAHEPF-----AKNDFITGESYA 229
QY 232 GHYGPAPFNHFYQNERIANGSVNGVQLNFNLSLGIINGIIDEATQAPYEFVANNYGI 291
Db 230 GHYIPAFARVHRGNK-----AKEGIHINLKGFAIGNGLTDPALQYKAYTDYALDGMV 284
QY 292 KAVNETVYNYMKFANQMPGQDILSTCKQTNR-TALADYALCAEATMCRDNVEGYPYA 350
Db 285 KSDHRI-----NKLVPVCEMAIKLCGTGDTTISCMASYFVC-----NNIFNGMA 329
QY 351 FAG-RGVYDIRHPYDDPTPPSYN--KFLAKDSYMDALGV-NINTQSNNDVYAFQQTG 406
Db 330 LAGDTNYVDVRKVCBGTLCYXD---NFTEDLEEILALPVRVSLIY- 384
QY 407 DFVWFNFIEDLEEILALPVRVS-----LIY-GDADYICNWFQGGQAVSLAANYSOAAQFRS 460
Db 385 --MLVDWMNLE--VGIPVLLEDGIKLLAVYAGEYDLICNWLGNRSRWHAMENSGQKEFVA 440
QY 461 AGYTPKVNVEYGETREYNFSFTRVYEAHEVYPIQPIASLQFNRTIFGWDTAEGQK 518
Db 441 SPEVPIVDGAEAGVLKTHGALGFLKVDHAGHVPMDQPKAALEMLKRWTKGTLSADS 500
QY 519 QKKI 522
Db 501 EKLIV 504

RESULT 7
Q9XH61 PRELIMINARY; PRT; 501 AA.
AC Q9XH61;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Serine carboxypeptidase.
OS Matricaria chamomilla.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae;
OC Anthemideae; Matricaria.
OX NCBI_TaxID=98504;
[1]
RN SEQUENCE FROM N.A.
RP Kohchi C., Yasuda H., Hirata T.;
RT "Isolation of a cDNA encoding for a carboxypeptidase, having leucine
RT zipper structure at the N-terminal region, from the cultured shoot
RT primordia of Matricaria chamomilla."
RL Plant Biotechnol. 16:409-412(1999).
RN [2]
RN SEQUENCE FROM N.A.
RP Kohchi C., Yasuda H., Hirata T.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141384; AAD42963.2; -
DR HSSP; P00729; IYSC.
DR MEROPS; S10.009; -
DR InterPro; IPR001563; Serine_carbpept.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PRODOM; PD001189; Serine_carbpept; 1.

DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 DR PROSITE; PS00131; CARBOXYPEPT SER; 1.
 KW Carboxypeptidase.
 SQ SEQUENCE 501 AA; 55973 MW; DBDFF8DB41980A66 CRC64;

Query Match 15.9%; Score 473.5; DB 10; Length 501;
 Best Local Similarity 27.8%; Pred. No. 5e-25;
 Matches 150; Conservative 83; Mismatches 213; Indels 93; Gaps 20;

QY 6 FLSVLPVLAASWALP-----GSTPAGV---GRQLPKNPTGV---KTLTTANN----- 47
 DB 8 FLSLLILAQSSSLPLQTLPTLSTQASSLLRGLNLPFKHVNIVEEDDVRNNSLDKKI 67
 QY 48 VTIKYKPG---ABVCETTEGVKSGY---VDTSPESHTEFWFFFAHNPETAPITLWL 102
 DB 68 VEKLFREFNFVESDGV---TVEDLGHAGYQIQHSHAAKMFYFFFSRNKKD-PVWIL 124
 QY 103 NGFGPSLGLPEELGCHVNSTFDDYINPHSWNEVSNLLFLSQPLGVGFYSYDVTGVS 162
 DB 125 TGGPGCSSELALFYENGFPIADNMTLVNEXGWDQASNLIVDQPTGTGFGFSYSDKRDI 184
 QY 163 INPVTGVVENSFAGVQGRYPTIDATLIDTNLAEEAAWEILQGLPSLDSRVQSKD 222
 DB 185 RHDEQGV-----SDDLDFLQAFTEPEF-----VDND 213
 QY 223 FSLWTESYGHHYGPAPFNHFEYONERIANGVNGVQNFNSLGIINGIIDEAIAQAPYYPE 282
 DB 214 FYITGESYAGHYPTAARVHOGNK-----AKEGIHNLKFGIGNGLTDPALQYQAYTD 268
 QY 283 FAVNNTYGIKAVNETVYNYMKEANQMGQODLISTCKQTNR-TALADYALCAEATNMCR 341
 DB 269 YALD-----MGLIKESQY---KRNLVPLCEAAIKLCGTGDTGVSCMAAYVVC----- 313
 QY 342 DNVGPPYAFAGR-GVYDIRHP-----YDDETPPSYNNKFLAKDSVMDAIGV-NINYTQ 393
 DB 314 NTIFSSILSTAGNINYYDIRKECVTSMCYDF-----SDMETLNKKSVRQALGVGDIIEVS 369
 QY 394 SNNDVYAFQOQDGFVWPNFIEDLEHILALPVRSVLIYGDYICNFGGQAVSLAANS 453
 DB 370 CSTVTYVYAMLVD---LMRNLEAGIPLELDGKMLVYAGEYDVCNWLGNRWVHAMEWN 426
 QY 454 QAAQFRSAGYTPUKVNGVEYGETREYGNFSFRVVEAGHEVYVYQPIASLQLFNRTIFG 512
 DB 427 GKQFNALSEAPEVDGSEAGLLKSGPLFLKVHDAGHWPMDQPKAALEMLKRWMDG 485

RESULT 8
 Q8GVTL ID Q8GVTL PRELIMINARY; PRT; 429 AA.
 AC Q8GVTL;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Carboxypeptidase C cbp31.
 GN P043A03.7.
 OC Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
 clone:P043A03."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF004299; BAC45113.1; --
 SQ SEQUENCE 429 AA; 47790 MW; 1F2F64F236475BB4 CRC64;

Query Match 15.9%; Score 473; DB 10; Length 429;
 Best Local Similarity 28.6%; Pred. No. 4.4e-25;
 Matches 130; Conservative 77; Mismatches 178; Indels 70; Gaps 16;

QY 70 YSGY--VDTSPESHTEFWFFFAHNP-ETAPITLWNGGSGSLGLPEELGCHVNST 122
 DB 17 HAGYRLPNTHDARLFYFFFSRSGSGEDDPVIVLWLTGGPGCSSELALFYENGFPIADN 76
 QY 127 FDDYINPHSWNEVSNLLFLSQPLGVGFYSYDVTGDSINPVTGVVENSFAGVQGRYPTID 186
 DB 77 MSLWVNDFGWDQSNLIYVDQPTGTGFGFSYSSNPRDTRHDEAG-VSNDLYA----- 125
 QY 187 ATLIDTNLAEEAAWEILQGLPSLDSRVQSKDPSLWTESYGGHYGPAPFNHFEYON 246
 DB 126 -----FLQAFTEPEHNF-----AKNDFYITGESYAGHYPTAFASRVYKGN 165
 QY 247 ERIANGSVNGVQNFNSLGIINGIIDEAIAQAPYEFPAVNTYGIKAVNETVYNYMKEAN 306
 DB 166 K-----NSEGIHNLKFGAIGNGLTDPALQYKAYTDYSLD-----MGLITKSFNRI--N 213
 QY 307 QMPNGCODLISTCKQTNR-TALADYALCAEATNMCRDNVEGVPYAFAGVGVYDIRHP--- 362
 DB 214 KIVPTCELAIKLGTSGTISCLGAYVVC-----NLFSSEI---TIIGKNYVDIRKPCVG 266
 QY 363 ---YDDETPPSYNNKFLAKDSVMDAIGV-NINYTQSNNDVYAFQOQDGFVWPNFIE-DL 417
 DB 267 SLCYD-----LSNMKEFLQKSVRESLGVGDIQFVSCPTVYQAML---LDWMRLNLEVG 318
 QY 418 EEITALPVRSVLIYGDYICNFGGQAVSLAANSQAAQFRSAGYTPUKVNGVEYGETR 477
 DB 319 PELLENIDKILVYAGYEDLICNWLGNRWVNSWMSGKEAFVSSSEEPFTVDGKEAGILK 378
 QY 478 EYGNFSFRVVEAGHEVYVYQPIASLQLFNRTIFG 512
 DB 379 SYGFLSFLKVHDAGHWPMDQPKVLEMLRWTS 413

RESULT 9
 Q9FFBO ID Q9FFBO PRELIMINARY; PRT; 505 AA.
 AC Q9FFBO;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Serine carboxypeptidase.
 OC Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=97471969; PubMed=9330910;
 RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
 features of the 1.6 Mb regions covered by twenty physically assigned
 P1 clones."
 RL DNA Res. 4:215-230(1997).
 DR EMBL; AB005243; BAB10619.1; --
 DR HSSP; P10619; IIVY.
 DR MEROPS; S10.009; --
 DR InterPro; IPR001563; Serine carboxpept.
 DR InterPro; IPR000379; Ser estrs site.
 DR Pfam; PF00450; serine carboxpept; 1.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine carboxpept; 1.
 DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 DR PROSITE; PS00131; CARBOXYPEPT SER; 1.
 KW Carboxypeptidase.
 SQ SEQUENCE 505 AA; 56543 MW; 6E74CD35109B099 CRC64;

Query Match 15.7%; Score 466.5; DB 10; Length 505;
 Best Local Similarity 28.3%; Pred. No. 1.6e-24;
 Matches 132; Conservative 75; Mismatches 178; Indels 81; Gaps 15;

QY 65 PCKVKS---YSGY---VDTSPESHPTFFWFPEARHNPTATITLWNGPGSDSLIGLFEELG 119
 Db 91 PSVQFEGHAGYSLPHSKSAKMFYFFESR-NKTTDPVILWLTGGPGSSVAMFYENG 149
 QY 120 PCHVNSTDDYINPHSWNEVSNLLFLSQPLGVGFSDTVGDSINPVTGVWENSFAGVQ 179
 Db 150 PFKISKDLSLWYNDGWDKVSNIIVDQPVGTGFSYTSDESIRNDEGVDV----- 199
 QY 180 GRYPTIDATLIDTTLNLAEEAEIILQGLSLPSLDSRVQSKDFSLWTSYGGHYGPAPF 239
 Db 200 -----SNDLDFLOAFKEHP-----KEVKNDFFITGESYAGHYPALA 238
 QY 240 NHFYEQNERIANGSVNGVQNFNSLGIINGIIDEAIQAPYEPFAVNNYTGKAVNETVY 299
 Db 239 SRVHSGNKK-----XEGIPINLKGFAIGNLGNLTPETQYGAIGDYALQ-----MKLISESDH 289
 QY 300 NYMKFANQWPGSCODLSTCKQTNRTALADYALCAEATNMCNCRDNVEGPPYAFAGRVYDI 359
 Db 290 ESLK-----QDYVEQNTITKCSGLGGLVCDGSAVEVCTSIKVIKXSGNLNYDI 340
 QY 360 RHP-----YDQTPPSYNNKFLAKDSVMDAIGV-NINTQSNNDVYVAFQQTGDFWPN 412
 Db 341 RKCVGSLCYDF-----SRMEIFLNKENVRKALGVGDIKFVSCSSTVY-----DAMIED 389
 QY 413 FIEDLEIILALPVRS-----LIY-CDAYICNWFQGVQAVSLAANYQAAQFRSAGYTPL 466
 Db 390 WMQNLB--VKIPSLVNDGINLLVYAGEYDLICNLGNLSRWVDQMNWSGQKFGSGAKNVSF 447
 QY 467 KNGVEYGETREYNFSTFRVYAGHEVYVYQPIASLQLENRTIFG 512
 Db 448 LVDGKEAGLLKNGHGLTLFKVYNAGHVMVMDQPKASLOMLQNNMQG 493

RESULT 10

Q8L6A7 PRELIMINARY; PRT; 508 AA.
 ID Q8L6A7
 AC Q8L6A7
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Carboxypeptidase type III.
 GN CP-III.
 OS Theobroma cacao (Cocoa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Malvales; Malvaceae; Byttnerioideae; Theobroma.
 OX NCBI_TaxID=3641;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tazi H., McCarthy J., Bucheli P., Laloi M.;
 RT "Molecular characterization of a type-III carboxypeptidase (cacao CP-
 III) from Theobroma cacao seeds.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ31314; CAC86383.1; -
 DR InterPro; IPR001563; Serine_carbpept.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR Pfam; PF00450; serine_carbpept; 1.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine_carbpept; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 SQ SEQUENCE 508 AA; 56521 MW; 68B63707DC686BFE CRC64;
 Query Match 15.6%; Score 464.5; DE 10; Length 508;
 Best Local Similarity 27.2%; Pred. No. 2.2e-24;
 Matches 153; Conservative 92; Mismatches 210; Indels 107; Gaps 23;
 QY 4 YEF-LSVLPIVAASWA-----LPGSTPASPVRQL-----PKNPTGKTLTTA 45
 Db 8 YPFGVSLFLTISISSAAASSFLDERRLGGSSFPISHAKLRIELNLFPEKEVNV---VDG 64
 QY 46 NNVTI-----RYKEPGAEGVCETTPGVKS-----YSGY---VDTSPESHPTFFWFFE 88

Db 65 QGVSLPESRLVERKFKFENL-----AVPGGVSVEDLGHAGYKLANSHDARMFYFRPE 119
 QY 89 ARHNPTETATITLWNGPGSDSLIGLFEELGCHVNSTDDYINPHSWNEVSNLLFLSQPL 148
 Db 120 SR-NSKQDPVILWLTGGPGCSSELALFYENGPFITAEANMSLIWNOYGMDSNLLYVDQ 178
 QY 149 LGVGSYSDDTVGDSINPVTGVWENSFAGVQGRYPTIDATLIDTTLNLAEEAEIILQGL 208
 Db 179 IGTGFSYS-----SDRRDIRNEDEVSNLDYD-----FLQAF 211
 QY 209 SGLPSLDSRVQSKDFSLWTSYGGHYGPAPFNFHFEQNERIANGSVNGVQNFNSLGIIN 268
 Db 212 AEHPF-----EKNDYIITGESYAGHYIIFAFARVHQGNK-----AKDGHINLKGFAIGN 262
 QY 269 GIIDEAIQAPYEPFAVNNYTGKAVNETVNYMKFANQWPGSCODLSTCKQTNRTAL 327
 Db 263 GLTDPAIQYKAYTDYALD-----MGVIKKSVDYNI---NKLVPVCMAIKLCGTGDTISCM 315
 QY 328 ADYALCAEATNMCNCRDNVEGPPYAFAG-RGVYDIRHPYDDPTPPSYN--KELAKDSVMDA 384
 Db 316 ASYFVC-----NAIFGTGINALAGDTNYDIRKCGSLCYDFSNNMETFLNQSVRDA 367
 QY 385 IGV-NINTQSNNDVYVAFQQTGDFWPNFIE-DLEEILALPVRVSLIYGDADYICNWF 442
 Db 368 LGVGSIDFVSCSPTVYQAMLV-----WNRNLEVGIIPALLEDGVKLLVYAGEYDLICNWL 423
 QY 443 GOAVSLAANYQAAQFRSAGYTPLKNGVEYGETREYNFSTFRVYAGHEVYVYQPIAS 502
 Db 424 NSRWVHAMEWSGQKFBFVASPEVPFVVDGSEAGVLRTHGFLGLKVDHAGHVMVMDQPKAA 483
 QY 503 LQLFNRTIIFG--DIAEGOKKI 522
 Db 484 LEMLKRWTKGTLSAADSEKLV 505

RESULT 11
 Q9LXH4 PRELIMINARY; PRT; 510 AA.
 ID Q9LXH4
 AC Q9LXH4
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Carboxypeptidase precursor-like protein (AT3G45010/F14D17_80).
 GN F14D17_80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jordan N., Bangert S., Wiedelmann R., Voss H., Unsel M., Mewes H.W.,
 RA Rudd S., Lemcke K., Mayer K.P.X., Quetier F., Salanoubat M.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Piam P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Arabidopsis cDNA clones";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.


```
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL353992; CAB89316.1; -
DR EMBL: AY050427; AAK91443.1; -
DR EMBL: AY033993; AAK16254.1; -
DR HSP: P00729; IYSC.
DR MEROPS: S10.009; -
DR InterPro: IPR001563; Serine carboxypept.
DR InterPro: IPR000379; Ser esterase.
DR Pfam: PF00450; serine carboxypept.
DR PRINTS: PR00724; CRBOXYPTASEC.
DR ProDom: PD001189; Serine carboxypept.
DR PROSITE: PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE: PS00131; CARBOXYPEPT_SER_SER; 1.
SQ SEQUENCE 510 AA; 56896 MW; 134BC30F4E64BF0A CRC64;

Query Match 14.7%; Score 439; DB 10; Length 510;
Best Local Similarity 26.6%; Pred. No. 1.4e-22;
Matches 123; Conservative 81; Mismatches 189; Indels 69; Gaps 14;

QY 64 TPGVKSY---SGV--VDTPESHTFFFEARHNPETAPITLWNGPGSDSLIGLPEEL 118
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
93 SPVSQDFGHAGYKYNFNSKAARMFYFFESRTN-KADPVVIMLTGGPGCSSELALFYEN 151
QY 119 GPCVNSTFDDYINPHSWNEVSNLLFLSQPLGVGFYSYSDTVGSIINPVTGVNENSPAGV 178
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
152 GPTVSNSSLSWNEFCWKNLIYVDQVGTGFGSYTSQSDLRHDEGCV----- 202
QY 179 QGRYPTIDATLIDTTNLAABAAWEILQGLFSLDSRVQSKDFSLWTSYSGHYGPAP 238
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
203 -----SNDLYDFLQAFKHEP-----QFVKNDFYITGESYAGHYIPAL 240
QY 239 FNFHYEQNERIANGSVNGVQLNFNSLGIINGIIDEAIOAPYPEFAVNNYTGKAVNETV 298
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 ASRVHRGNK-----NKEGTHINLKGFAIGNGLNPEIQGAYADYALDWNMLITQSDHNL 295
QY 299 YNYMKEANQMPNGCQDLISTCKQTNRTALADYA---LCAEATNMCNDNVEGPPYAFAGR- 354
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 NRY-----YATCQSIKECSADGEGDACCASSYTVCCNNIFQKIMDIAGNV 340
QY 355 GVYDIRHPYDDPTPPSYN--KFLAKDSVMDAIGV--NINYTQSNNDVYAFQOTGDFVWP 411
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
341 NYDYVRKQCEGSLCYDFSNNENFLNQSVRKALGVGDIEFVSCSTAVYEAQMD----WM 396
QY 412 NFIE-DLEBILALPVRVSLIYGDAVYICNWFQGVQAVSLAANYSOAQAQFSAGYTPKLVNG 470
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
397 RNLEVGIPALLQDGIGKLLVYAGEYDILCNWLGSKWVHEMWSGQKEFVAAATVPFHVGN 456
QY 471 VYGETREYGNFSFTRVYAGHEVPYQPIASQLFNRTIFG 512
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
457 KEAGLMKNYGSITFLKVDHAGHVPMDQPKAALQMLQNMVQ 498

RESULT 12
Q93ZC3
ID Q93ZC3 PRELIMINARY; PRT; 510 AA.
AC Q93ZC3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE AT3945010/F14D17.80.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY057639; AAL15270.1; -
DR InterPro: IPR001563; Serine carboxypept.
DR InterPro: IPR000379; Ser esterase.
DR Pfam: PF00450; serine carboxypept.
DR PRINTS: PR00724; CRBOXYPTASEC.
DR ProDom: PD001189; Serine carboxypept.
DR PROSITE: PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE: PS00131; CARBOXYPEPT_SER_SER; 1.
SQ SEQUENCE 510 AA; 56837 MW; 704BC00C4E64BF08 CRC64;

Query Match 14.7%; Score 438; DB 10; Length 510;
Best Local Similarity 26.6%; Pred. No. 1.6e-22;
Matches 123; Conservative 80; Mismatches 190; Indels 69; Gaps 14;

QY 64 TPGVKSY---SGV--VDTPESHTFFFEARHNPETAPITLWNGPGSDSLIGLPEEL 118
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
93 SPVSQDFGHAGYKYNFNSKAARMFYFFESRTN-KADPVVIMLTGGPGCSSELALFYEN 151
QY 119 GPCVNSTFDDYINPHSWNEVSNLLFLSQPLGVGFYSYSDTVGSIINPVTGVNENSPAGV 178
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
152 GPTVSNSSLSWNEFCWKNLIYVDQVGTGFGSYTSQSDLRHDEGCV----- 202
QY 179 QGRYPTIDATLIDTTNLAABAAWEILQGLFSLDSRVQSKDFSLWTSYSGHYGPAP 238
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
203 -----SNDLYDFLQAFKHEP-----QFVKNDFYITGESYAGHYIPAL 240
QY 239 FNFHYEQNERIANGSVNGVQLNFNSLGIINGIIDEAIOAPYPEFAVNNYTGKAVNETV 298
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 ASRVHRGNK-----NKEGTHINLKGFAIGNGLNPEIQGAYADYALDWNMLITQSDHNL 295
QY 299 YNYMKEANQMPNGCQDLISTCKQTNRTALADYA---LCAEATNMCNDNVEGPPYAFAGR- 354
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 NRY-----YATCQSIKECSADGEGDACCASSYTVCCNNIFQKIMDIAGNV 340
QY 355 GVYDIRHPYDDPTPPSYN--KFLAKDSVMDAIGV--NINYTQSNNDVYAFQOTGDFVWP 411
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
341 NYDYVRKQCEGSLCYDFSNNENFLNQSVRKALGVGDIEFVSCSTAVYEAQMD----WM 396
QY 412 NFIE-DLEBILALPVRVSLIYGDAVYICNWFQGVQAVSLAANYSOAQAQFSAGYTPKLVNG 470
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
397 RNLEVGIPALLQDGIGKLLVYAGEYDILCNWLGSKWVHEMWSGQKEFVAAATVPFHVGN 456
QY 471 VYGETREYGNFSFTRVYAGHEVPYQPIASQLFNRTIFG 512
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
457 KEAGLMKNYGSITFLKVDHAGHVPMDQPKAALQMLQNMVQ 498

RESULT 13
Q17679
ID Q17679 PRELIMINARY; PRT; 2105 AA.
AC Q17679
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Y1684A.2 protein.
GN Y1684A.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
```

RP SEQUENCE FROM N.A.
RA Thomas K.
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z70203; CAA94110.1; -. JOINED.
DR EMBL; Z70203; CAA94110.1; JOINED.
DR EMBL; Z93339; CAA94110.1; JOINED.
DR EMBL; Z93339; CAB07544.1; -.
DR EMBL; Z93339; CAB07544.1; JOINED.
DR EMBL; Z70203; CAB07544.1; JOINED.
DR EMBL; Z70203; CAB07544.1; JOINED.
DR EMBL; Z93339; CAA19443.1; JOINED.
DR EMBL; Z70203; CAA19443.1; JOINED.
DR HSP; P10619; 11VY.
DR WormPep; Y16B4A.2; CB211374.
DR InterPro; IPR001563; Serine carbpept.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine carbpept; 4.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PROSITE; PS001189; Serine carbpept; 6.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 4.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 2.
SQ SEQUENCE 2105 AA; 234335 MW; 139CE6E96D75B3FB CRC64;

Query Match 14.0%; Score 417.5; DB 5; Length 2105;
Best Local Similarity 23.5%; Pred. No. 3.6e-20;
Matches 156; Conservative 91; Mismatches 209; Indels 209; Gaps 28;

QY 2 RYEFSLVPLVAASWA-----LPGSTPASVGRQLPKNPTGVKTLTANN 47
DB 1467 QSQVYTTIAGKSWTQNLVLTWKGSHFVSDRPAQ-----ALQMLNFI 1515
QY 48 VTIYKEPAGV-----CET-----TPGV-----KSYGVVDTS 80
DB 1516 NOANSTPAGIDVTPQPLTNVTSQNTCTGTQDRIINLPLGADMLFKCYSGFLDGLSGH 1575
QY 81 HTFFWFEARHNPETAPITLWNGGSDSLGLFEELGFCCHVNSTFDYI-NPHSWNEV 139
DB 1576 KVHYLVSENNPSTPDLILLWNGGSDSLGLFEELGFCCHVNSTFDYI-NPHSWNEV 1635
QY 140 SNLFLSQPLGVGFSYDVTGDSINPTGVNENSSPAGVQGYPTIDATLDTNLAEEA 199
DB 1636 ANVLYESPIGVGYSA-----YNTNIQYDDVT--TAG 1668
QY 200 AMEILQGLSLGSLDSRVQKDFSLWTSYGYGHYPAFFNFHYEQMERIANGSVNG-VQ 258
DB 1669 NYAALKSFFAQP---QYTTSDFYTTGESYAGVYLPGL-----SALLVQGIKGDIN 1717
QY 259 LNFNSLGIINGIDE-----ALQAPYPEFAVNNYTGKAVNETVYN----- 300
DB 1718 INYKGSIGNGVIDKRTDMNSQLHYQY-----HG--GISATTYNTALQLCCSGDEF 1767
QY 301 -----YMKFANQMPNG-----CQD-LISTCKQTKRNTALADYALCAEATNCRDNVEGP 347
DB 1768 KCFSDRMNFNNSIFWGLSDSPCYDFVATGALLTAFFDPYMWYQOCWTI----- 1819
QY 348 YZAFAGRGVYDIRHPYDDPTPPSYNKK-----FLAKDSV 381
DB 1820 -----PYNDITPRPYGETWTGTGINYSSDFLNGPYCDDSGAMEGLNRPVV 1865
QY 382 MDAIGV--NINTQSNNDVYFAQQGDFWPNFIEDLEILA---LPREVSILYGDY 436
DB 1866 RKALNIPDSYPYWAANNIINAYNQVDSIVPN---LQIIMANAPANFKMLLYSGDADT 1921
QY 437 ICNWFQGGQ-----AVSLAANYSOAA---QFRSAGYTPKLVNGVEYGETREYGNFSFT 485
DB 1922 MNWLGAETPANNFAALGLTSSPRAQWTFQYNST-FQP-TVAGYQTSYTNAINIDLVL 1979
QY 486 RYVEAGHEVYVYOFIASLQFNRTIFQWDIAGOKKIWPYKXNGTATATHQSSVP--L 543
DB 1980 TVKSGSHFVLDLPQALQW-----IYNFVKSRGVNTFPDLNSFTTTTTTTTTPGTG 2035
QY 544 PTATS 548

DB 2036 PTVTA 2040

RESULT 14
Q9FMX9 PRELIMINARY; PRT; 499 AA.
AC Q9FMX9;
DC 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serine carboxypeptidase II-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones.";
RL DNA Res. 4:401-414(1997).
DR EMBL; AB007648; BAB11176.1; -.
DR HSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR001563; Serine carbpept.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PROSITE; PS001189; Serine carbpept; 2.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 499 AA; 56329 MW; 6150FB91A2E7BCBD CRC64;

Query Match 13.8%; Score 411.5; DB 10; Length 499;
Best Local Similarity 26.9%; Pred. No. 1.1e-20;
Matches 152; Conservative 80; Mismatches 216; Indels 117; Gaps 26;

QY 5 EFLSVLPLVAASWALPGSTPASVGRQLPKNTG-----VKLTITANNVTIRKE-PAEG 59
DB 7 EF-SVLVPLVLSFLL--GSTA-----EKLCSNDGNGCFRSVLAQAQADRVKELFG--- 57
QY 60 VCETTPGVK--SYSGVVDTSPESSH---TFWFEEARHNPETAPITLWNGGSDSL-IG 113
DB 58 ----QPVKFRQVAGIV-TVNETHGALFYWFEATQNSKPEVLWLNNGGPGCSIGFG 112
QY 114 LFEELGFCCHVNSTFDY--YINPHSWNEVSNLLFLSQPLGVGFSYSTVPGSINPVTGWVE 171
DB 113 AAELGPFPPQNSQPKLNFYSNMKAAALLFLESFVGFGFSYNT----- 159
QY 172 NSSFAGVQGYPTIDATLDTNLAEEAAWEILQGLSLGSLDSRVQKDFSLWTSYNG 231
DB 160 -----SRDIKQLGDT-VTARDSYNFLNWNKFRFP-----QYKSHDFVIAGESVA 202
QY 232 GHYGAFFNFHYEQMERIANGSVNGVQLNFSLGIINGIDEALQAPYYPEFAVNNYTG 291
DB 203 GHVTPQGLSELIYKENKIAKKDF-----INLKLMIKNALLDDTDQKMEYAWDHAVIS 258
QY 292 KAVNETVYNYMKFANQM--PNGQODLIS-----TCKQNTKRALADYALC 333
DB 259 DALYKVNKNKCDFKQKLVTKECNDALDEYDPVYKILDMYSLYAPKCVPTSTNSSTSHVA 318
QY 334 AEATNCRDNVEGYPYAFAGRGVYDIRHPYDDPTPPSYNKKFLAKDSVMDAIGVNYTQ 393
DB 319 GNRPLPAFRSILPRILISNEGWRMAAGY-DPCASEYTEKYNKDKVQBALHNV----- 373
QY 394 SNNDVYVAFQQTGDFV--W-----PNFTEDLEEILALPVRVSLYIGDADYICNMFQGA 447

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:54:40 ; Search time 17 Seconds
(without alignments)
1535.284 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 2979

Sequence: 1 MRGYEFLSLVLAASWALP.....HTQSSVPLPLTATSSSVGMA 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	894	30.0	423	1	CPS1_PENJA
2	652.5	21.9	1002	1	CBPY_SCHPO
3	636	20.1	508	1	YBY9_YEAST
4	599	20.1	542	1	CBPY_CANAL
5	594.5	20.0	523	1	CBPY_PICPA
6	527.5	17.7	532	1	CBPY_YEAST
7	475	15.9	516	1	CBPX_ARATH
8	473	15.9	429	1	CBPX_ORISA
9	455	15.3	500	1	CBP3_ORISA
10	437	14.7	508	1	CBP3_HORVU
11	416.5	14.0	500	1	CBP3_WHEAT
12	359	12.1	476	1	CPVL_HUMAN
13	355.5	11.9	470	1	YSS2_CAEEL
14	353.5	11.9	436	1	CP22_HORVU
15	351	11.8	482	1	NF31_NAEFO
16	349	11.7	474	1	PRTP_MOUSE
17	347.5	11.7	516	1	CP23_HORVU
18	333	11.2	480	1	PRTP_HUMAN
19	332	11.1	469	1	YUM5_CAEEL
20	331	11.1	729	1	KEK1_YEAST
21	326.5	11.0	454	1	YUM6_CAEEL
22	318.5	10.7	476	1	CBP2_HORVU
23	316	10.6	574	1	YBP3_CAEEL
24	312.5	10.5	471	1	VCP_AEPAE
25	312	10.5	510	1	CBP1_ORISA
26	305.5	10.3	423	1	CBP2_WHEAT
27	297.5	10.0	505	1	YX02_CAEEL
28	295.5	9.9	286	1	CBPX_PEA
29	280	9.4	499	1	CBP1_HORVU
30	271	9.1	523	1	PEPS_ASPSA
31	255	8.6	452	1	RISC_MOUSE
32	248	8.3	452	1	RISC_RAT
33	237	8.0	452	1	RISC_HUMAN

34	221.5	7.4	531	1	PEPF_ASNG
35	219.5	7.4	507	1	SXA2_SCHPO
36	167.5	5.6	324	1	CP21_HORVU
37	147	4.9	366	1	HNLS_SORBI
38	129.5	4.3	1061	1	OAR_MYXA
39	114.5	3.8	717	1	AMVM_BACST
40	113.5	3.8	1287	1	RPO1_FOWPV
41	112	3.8	376	1	OE56_NPVAC
42	112	3.8	787	1	AGL2_BACTQ
43	111	3.7	696	1	OXLA_NEUCR
44	110.5	3.7	1196	1	AMVB_PABPO
45	109.5	3.7	481	1	PRTB_ERWCH

ALIGNMENTS

RESULT 1

CPS1_PENJA STANDARD; PRT; 423 AA.
AC P34946;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase S1 (EC 3.4.16.6)
OS Penicillium janthinellum (Penicillium vitale).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5079;
RN [1]
RP MEDLINE=94039747; PubMed=8224168;
RX Svendsen I., Hofmann T., Endrizzi J., Remington S.J., Breddam K.;
RA "The primary structure of carboxypeptidase S1 from Penicillium
janthinellum";
RL FEBS Lett. 333:39-43(1993).
CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
or lysine residue.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
DR PIR; S38953; S38953.
DR HSSP; P00729; LYSC.
DR MEROPS; S10.008; -.
DR InterPro; IPR000379; Ser_estrs_site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein.
FT ACT SITE 143 143 POTENTIAL.
FT ACT SITE 340 340 BY SIMILARITY.
FT ACT SITE 397 397 BY SIMILARITY.
FT BINDING 343 343 SUBSTRATE.
FT BINDING 398 398 SUBSTRATE.
FT DISULFID 8 68
FT DISULFID 55 300
FT DISULFID 223 246
FT DISULFID 230 239
FT CARBOHYD 200 200 N-LINKED (GLCNAC...).
SQ SEQUENCE 423 AA; 46496 MW; 55F3333608B12729 CRC64;

Query Match 30.0%; Score 894; DB 1; Length 423;
Best Local Similarity 40.1%; Pred. No. 5.7e-55;
Matches 186; Conservative 66; Mismatches 160; Indels 52; Gaps 8;

QY	59	GYCETTPGVKSGYVDSPESHTFWFPEARHNDETAPITLWLNKGPDSGLIGLFEEL 118
Db	6	GICETTPGVKSGYVDSPESHTFWFPEARHNDETAPITLWLNKGPDSGLIGLFEEL 65
QY	119	GFCH-VNSTFDDYINPHSWNEVSNLLELSQPLGVGFYSYSDTVGDSINPVTGVWENSPAG 177
Db	66	GFCHVNGSDTSPSLNENSWNNYNNIYIDQPIGVGFYSY----- 104

Db	71	LDGLTPEIKNIWSEMLKFP	-----NSGITEINFKAPKKCKKITQQDFPHVTDQV	121	
Qy	64	-----TP-----	GVKSYSGYVDTSPE-SHTFFWFFFAARNPETATITWLNGGPG	107	
Db	122	PNHKLRIKSTPKDLGIDTVKQSGYGLDVVDKDKHFFYFFESRNDPKNDPVLWNGGPG	181		
Qy	108	SDSLIGLFEELGCPCHVNSTFDDYINPHSWNEVSNLLFLSQPLGWGFSYSDVDGGINPVT	167		
Db	182	CSLSLGLFELGPGSSIDKMLKPVNPHSWNANASVIFLDQPINVGISYS	230		
Qy	168	GVVENSFPAGVQGRYPTDAILDTITNLAAEAAWEILQFLSLGFLSDRSRVQSKDFSLWT	227		
Db	231	-----SQSVNTIAAGKDVVAFQLQFFKNFP	---EVANLDFHIAG	267	
Qy	228	ESYGHYGPAPFNHNYQNERIANGSVNGVOLNFNSLGI	---NGHIDEAIOAPYPPPEAV	285	
Db	268	ESYAGHYTPAPASEILTHFER	-----NFNLTSVLINGLTDPLVQVEYEPNMAC	316	
Qy	286	NNTYGIKAVETVYNYMKFANOMPNGCODLISTCKOTNKRLADYALCAEATNMCRDNVE	345		
Db	317	GEGGEPSPVLEPECTQM	---LNSLPR-CLSLIESCYES	---GSVWSCVPATIIYCNNGQM	368
Qy	346	GPYIAFAGRGVYDIRHPYDDP	---TPPSYNNKFLAKDSVMDAIGVNIIN	---TQSNNDVY	400
Db	369	GPYQK-TGRNVYDIRTMCBGGSLCYSOLEYIDQYLNLP	PEVKKALGAEDVEQSCNFDINR	427	
Qy	401	AFQQTGDFWP	---NFIEDLEHILALPVRSLIY	---GDADYICNWPFGGQAVLSAANYSOAA	456
Db	428	NFMFAGDMWPKPYQKNVIDLEK	---ELFPV	---LIYAGDKDFICNLWLNQNAVTRNLWSGSK	482
Qy	457	QFSAGYTPCLKVNGVEYETREYGNFSTFRVYEAAGHEVFPYQPIASLQLFNRTIFG	512		
Db	483	GFTKAPVSKWVGKNAAGEVKYKHFTFLRVGGGHVFPYDQENALDMVNRWISG	538		

RESULT 5

CBPV_PICPA

ID CBPV_PICPA STANDARD; PRT; 523 AA.

AC P527J0;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCV).

GN PRC1.

OS Pichia pastoris (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Pichia.

ON NCBI_Faxid=4922;

FX [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 108-142.

RC STRAIN=GTS115;

EX MEDLINE=96381245; PubMed=8789258;

EA Ohi H., Ohtani W., Okazaki N., Furuhashi N., Ohmura T.;

RT "Cloning and characterization of the Pichia pastoris PRC1 gene encoding carboxypeptidase Y.";

RL Yeast 121:31-40(1996).

CC -!- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.

CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a broad specificity.

CC -!- SUBCELLULAR LOCATION: LYCOSOME-LIKE VACUOLES.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X87987; CAA61240.1; -.

DR PIR; S61713; S61713.

Db 463 PTFLEIYDAGHWVFDQPEASLEMWNGWISG 494

RESULT 4
CBPY_CANAL STANDARD; PRT; 542 AA.
ID AC P30574;
DT 01-APR-1993 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
CPYL
OS Candida albicans (Yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=93051356; PubMed=1427093;
RA Mukhtar M., Logan D.A., Kaufner N.F.;
RT "The carboxypeptidase Y-encoding gene from *Candida albicans* and its
transcription during yeast-to-hyphae conversion.";
RL Gene 1211173-177(1992).
CC -1- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -1- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
CC -1- INDUCTION: TRANSIENTLY DOWN-REGULATED DURING THE EARLY EVENTS OF
YEAST TO HYPHAE CONVERSION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch).

CC EMBL; M95182; AAA34326.2; -
CC HSP; P00729; 1CPY.
CC MEROPS; S10.001; -
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine_carppept.
DR Pfam; PF00450; serine_carppept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PRODOM; PD001189; Serine_carppept; 1.
DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 21
FT PROPEP 22 127
FT CHAIN 128 542
FT ACT_SITE 269 269
FT ACT_SITE 461 461
FT ACT_SITE 518 518
FT BINDING 464 464
FT BINDING 519 519
FT DISULFID 182 421
FT DISULFID 316 330
FT DISULFID 340 363
FT DISULFID 347 356
FT DISULFID 385 391
FT CARBOHYD 213 213
FT CARBOHYD 291 291
FT SEQUENCE 542 AA; 61044 MW; 7FA6B3F82F9D44AF CRC64;

Query Match 20.1%; Score 599; DB 1; Length 542;
Best Local Similarity 31.5%; Pred. No. 2.6e-34;
Matches 169; Conservative 64; Mismatches 193; Indels 110; Gaps 20;

FT	114	STRAND	114
FT	118	HELI	116
FT	129	STRAND	129
FT	135	TURN	138
FT	139	STRAND	146
FT	152	TURN	154
FT	157	STRAND	161
FT	164	TURN	166
FT	167	STRAND	167
FT	169	TURN	169
FT	170	HELI	173
FT	177	TURN	177
FT	180	STRAND	183
FT	184	TURN	186
FT	187	STRAND	190
FT	192	TURN	193
FT	195	HELI	198
FT	200	STRAND	202
FT	208	TURN	209
FT	211	TURN	212
FT	214	STRAND	215
FT	224	HELI	224
FT	240	HELI	240

Query Match	1
Best Local Similarity	2
Matches 149; Conservat	3
	4
	5
	6
	7
	8
	9
	10
	11
	12
	13
	14
	15
	16
	17
	18
	19
	20
	21
	22
	23
	24
	25
	26
	27
	28
	29
	30
	31
	32
	33
	34
	35
	36
	37
	38
	39
	40
	41
	42
	43
	44
	45
	46
	47
	48
	49
	50
	51
	52
	53
	54
	55
	56
	57
	58
	59
	60
	61
	62
	63
	64
	65
	66
	67
	68
	69
	70
	71
	72
	73
	74
	75
	76
	77
	78
	79
	80
	81
	82
	83
	84
	85
	86
	87
	88
	89
	90
	91
	92
	93
	94
	95
	96
	97
	98
	99
	100
	101
	102
	103
	104
	105
	106
	107
	108
	109
	110
	111
	112
	113
	114
	115
	116
	117
	118
	119
	120
	121
	122
	123
	124
	125
	126
	127
	128
	129
	130
	131
	132
	133
	134
	135
	136
	137
	138
	139
	140
	141
	142
	143
	144
	145
	146
	147
	148
	149
	150
	151
	152
	153
	154
	155
	156
	157
	158
	159
	160
	161
	162
	163
	164
	165
	166
	167
	168
	169
	170
	171
	172
	173
	174
	175
	176
	177
	178
	179
	180
	181
	182
	183
	184
	185
	186
	187
	188
	189
	190
	191
	192
	193
	194
	195
	196
	197
	198
	199
	200
	201
	202
	203
	204
	205
	206
	207
	208
	209
	210
	211
	212
	213
	214
	215
	216
	217
	218
	219
	220
	221
	222
	223
	224
	225
	226
	227
	228
	229
	230
	231
	232
	233
	234
	235
	236
	237
	238
	239
	240
	241
	242
	243
	244
	245
	246
	247
	248
	249
	250
	251
	252
	253
	254
	255
	256
	257
	258
	259
	260
	261
	2

[illegible]

Db 326 ERLCGLIESCDSQ-----SWSCVPATYICNNAQLAPYQR-TGRNVYDIRKCEGGLNC 379

QY 366 -PTPPSYNNKFLAKSDVMDAIGVNNVNTQS--NNDVYVAFQGTGVFVNPFIEDLEILAL 423

Db 380 YPTLQD-IDDYLNQDYVKEAGVDEHYESCNPFINNFILFAGDWMP-VHTAVTDLLNQ 437

QY 424 PVRVSLIYGADYICNWFQGVQVSLAANYSOAAQFRS---AGYPLKVGNGVEYGETREYG 480

Db 438 DLPILVYAGDKDICNWLGNKAWTDVLPWKYDEEFASQKRVNWTASITDEVA-GEVKSXK 496

QY 481 NFGFTVRYEAGHEVPVYQPTIASLQLENRTIFG 512

Db 497 HFTYLVFNGCHVPPDPVNPALSMVNEWIHG 528

RESULT 7

CBPX ARATH STANDARD; PRT; 516 AA.

ID AC CBPX ARATH STANDARD; PRT; 516 AA.

DT 01-OCT-1993 (Rel. 27, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Serine carboxypeptidase precursor (EC 3.4.16.-)

GN AT3G10410 OR F13M14.32

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Bradley D.;

RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=21016720; PubMed=11130713;

RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseid M.,

RA Farmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,

RA Deleney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,

RA De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,

RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quettier F.,

RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,

RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,

RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,

RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,

RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,

RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,

RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Masuy D.,

RA Cooke R., Laudie M., Berger-Liauro C., Purnelle B., Vitale D.,

RA Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,

RA Monfort A., Argirou A., Flores M., Lignori R., Zaccaria P., Mewes H.-W.,

RA Mannhaupt G., Haase D., Schoof H., Rudd S., Tallon L.J., Jenkins J.,

RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Zaccaria P., Mewes H.-W.,

RA Rooney T., Rizzo M., Walters A., Utterback T., Fujii C.Y., Shea T.P.,

RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,

RA Pai G., Millscher J., Sellers P., Gill J.E., Feldblyum T.V.,

RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,

RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,

RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,

RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Nakayama S., Nakazaki N., Shimp S., Takeuchi C., Wada T.,

RA Watanabe A., Yamada M., Yasuda M., Tabata S.;

RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis

thaliana.";

RL Nature 408:820-822 (2000).

RN [3]

RP SEQUENCE OF 252-372 AND 455-516 FROM N.A.

RC STRAIN=cv. Columbia; TISSUE=Seedling;

RA Hofte H.;

RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a

broad specificity.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC

CC EMBL; M81130; AAB04606.1; -

DR EMBL; AC011560; AAG51389.1; -

DR EMBL; Z59555; CAA81115.1; -

DR EMBL; Z46528; CAA81299.1; -

DR HSSP; P00729; IYSC.

DR MEROPS; S10.009; -

DR InterPro; IPR000379; Ser esters site.

DR InterPro; IPR001563; Serine carboxpept.

DR Pfam; PF00450; serine carboxpept; 1.

DR PRINTS; PR00724; CRBOXYPTASEC.

DR ProDom; PD001189; Serine carboxpept; 1.

DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.

DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.

DR Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.

KW SIGNAL

FT PROPEP 1 22 POTENTIAL.

FT CHAIN 23 82 POTENTIAL.

FT ACT_SITE 83 516 SERINE CARBOXYPEPTIDASE.

FT ACT_SITE 229 229 BY SIMILARITY.

FT ACT_SITE 417 417 BY SIMILARITY.

FT ACT_SITE 474 474 BY SIMILARITY.

FT BINDING 420 420 SUBSTRATE (BY SIMILARITY).

FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).

FT DISULFID 139 379 BY SIMILARITY.

FT DISULFID 300 322 BY SIMILARITY.

FT DISULFID 307 315 BY SIMILARITY.

FT CONFLICT 515 516 QM -> ADVITSSPFLHKNKQOIIKQCVSN (IN REF.

FT

SQ SEQUENCE 516 AA; 57301 MW; 177C778DF657A1C1 CRC64;

Query Match 15.9%; Score 475; DB 1; Length 516;

Best Local Similarity 30.1%; Pred No 1e-25; Indels 80; Gaps 18;

Matches 139; Conservative 71; Mismatches 172;

QY 63 TTPGVKVSXGY--VDTSPESHFFWFPEARHNPETAPITLWNGPGSDSLIGLFEELGP 120

Db 93 TVDDLGHAGYKLPKRGASMFYFFESR-NKKDAPVIVLWLTGGGCSSELAVFENG 151

QY 121 CHVNSTEDDYINPHSWNEVSNILFSLQPLGVGFYSYDVGSDINPVTVGWENSSPAGVQG 180

Db 152 FKITSNMSLAWNEYGWDQVSNLLVYDQVGTGFSYTTDKSDIRHDETVG----- 200

QY 181 RYPTIDATLIDTTMLAAEAANEILLQGLSLPSLSDRSVQSKDFSLWTSYGHGYPAPEN 240

Db 201 -----SNDLYDFLQAFPAEPKPL-----AKNDFYITGESYAGHYIPAPAS 240

QY 241 HFYQNERIANGSVNGVQLNFSNGIINGIIDEAIQAPYEPFVAVNNTYGIKAVNNTVYN 300

Db 241 RVHKGK--AN---EGVHNLKGFAGNGLTDPALQVPAYPDYALE--MGLITQKE----- 289

QY 301 YMKFANQWMPNGCQDLISTCKQTNETA-LADYALCAEATNMCRDNVEGVYPAFAGRGVYDI 359

Db 290 HDRLEKIVPL-CEUSIKLCGTGDTTSLCLASLVNCSLFGVMSHAGGVNY-----YDI 341

QY 360 RHP-----YDDPTPPSYNNKFLAKSDVMDAIGV-NINYTQSNNDVYVAFQGTGVFVWN 412

Db 342 RKKCVGSLCYDF-----SNMEKFLNQSVRKSLGVGDIDFVSCSTSVYQA-----MLVD 390

QY 413 FIEDLEILALPVRS-----LIY-GDADYICNWFQGVQVSLAANYSOAAQFRSAGYTEL 466

Db 391 WMRNLE--VGIPITLEDIGSILVYAGEYDILCNWLNRSRWNAWENSGKTNFGAAKEVPF 448

QY 467 KVGVEGETREYGNFSFTVYAGHEVPPYQPTIASLQLENRTIFG 508

Db 449 IVDGKEAGLLKTYEQLSFLKVRDAGHMVPMQPKAALMKLR 490

RESULT 8
CBPX ORYSA STANDARD; PRT; 429 AA.
ID CBPX ORYSA
AC P52712;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase-like precursor (EC 3.4.16.-).
GN CBP31.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Yukihikari;
RA Washio K., Ishikawa K.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; D1587; BAA04511.1; -
DR PIR; T03607; T03607.
DR HSP; P00729; IYSC.
DR MEROPS; S10.009; -
DR Gramene; P52712; -
DR InterPro; IPR000379; Ser estrs site.
DR InterPro; IPR001563; Serine carboxpept.
DR Pfam; PF00450; serine_carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydroxylase; Carboxypeptidase; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN ? 429
FT ACT_SITE 148 148
FT ACT_SITE 336 336
FT ACT_SITE 393 393
FT BINDING 339 339
FT CARBOHYD 76 76
FT CARBOHYD 414 414
FT CARBOHYD 417 417
FT SEQUENCE 429 AA; 4774 MW; 1D5A68544325B31 CRC64;
Query Match 15.9%; Score 473; DB 1; Length 429;
Best Local Similarity 28.6%; Pred. No. 1.1e-25;
Matches 130; Conservative 77; Mismatches 178; Indels 70; Gaps 16;
QY 70 YSGY--VDTSPEHTEFWFPEARNP-ETAPITLWNGGSGSLGLFEELGCHVNST 126
Db 17 HAGYRLPNTHDARLFYFFEFESRGSKGDDPVWLTGGPGCSSELALFYENGPFPIADN 76
QY 127 FDIYINPHSNWVNSLFLQPLCGVRSYSDTVGDSINPVYGVNSSSFAGVGRVPTID 186
Db 77 MSLVNDFGMDQESNLNLYVDQPTGTGFSYNSNPRDTHDAG--VNSDLA----- 125
QY 187 ATLIDTINLAARAAWEILQGLSLGSLDSRVQSKDPSLWTESYGGHYGPAFFNFHFQON 246
Db 126 -----FLOAEFTTEHPNF---AKNDFYITGESYAGHYIYAFASRVYKGN 165

QY 247 ERIANGSVNQVQNFNSLGIINGIIDEAIOCAPYYPFAVNTYGIKAVNETVYNYMKFAN 306
Db 166 K-----NSEGIHNLKGFALNGLTDPAIQYKATYDLSL-----MGLITKSQFNRI--N 213
QY 307 QMPNGCQDLISTCKQTNR-TALADYALCABATNCRDNVEGPIYAFAGRGVYDIRHP--- 362
Db 214 KIVPTCELAIKLCTSGTISCLGAYVVC-----NLICSSIE---TIIGKNVYDIRKPCVG 266
QY 363 ---YDDPTPEPYNNKFLAKDSVMDAIGV-NINTQSNNDVYAFQQTGDFWPNFIE-DL 417
Db 267 SLCYD-----LSNMEKFLQLKVSRESLGVGDIQFVSCSPVYQAML-----LDWMENLEVGI 318
QY 418 EEIALPVRVSLYGDADYICNWFQGGQAVSLAANYSOAOFRSAGYTPKLVNGVEYGETR 477
Db 319 PELLENDIKVLIVAGEYDLICNWLGNRWNMSMEWSKEAFVSSSEEPFTVDGKEAGILK 378
QY 478 EYGNFSFTRVYEAGHEVYPIYOPTIASLQLFNRTIFG 512
Db 379 SYGSLFLKVDHAGHMVPMQPKVALEMLMWTSG 413

RESULT 9
CBP3 ORYSA STANDARD; PRT; 500 AA.
ID CBP3 ORYSA
AC P37891;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase III precursor (EC 3.4.16.5).
GN CBP3.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Yukihikari; TISSUE=Seed;
RX MEDLINE=92329723; PubMed=1627776;
RA Washio K., Ishikawa K.;
RT "Structure and expression during the germination of rice seeds of the
gene for a carboxypeptidase.";
RL Plant Mol. Biol. 19:631-640(1992).
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SUBUNIT: Monomer (Probable).
CC -!- INDUCTION: BY GIBBERELIC ACID (GA). INHIBITED BY ABSICISIC ACID
(ABA).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; D10985; BAA01757.1; -
DR PIR; S22530; S22530.
DR HSP; P00729; IYSC.
DR MEROPS; S10.009; -
DR Gramene; P37891; -
DR InterPro; IPR000379; Ser estrs site.
DR InterPro; IPR001563; Serine carboxpept.
DR Pfam; PF00450; serine_carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydroxylase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 21
FT CHAIN 1 21

```

FT PROPEP 22 73 BY SIMILARITY.
FT CHAIN 74 484 SERINE CARBOXYPEPTIDASE III.
FT PROPEP 485 500 BY SIMILARITY.
FT ACT SITE 216 216 BY SIMILARITY.
FT ACT SITE 404 404 BY SIMILARITY.
FT ACT SITE 461 461 BY SIMILARITY.
FT BINDING 407 407 SUBSTRATE (BY SIMILARITY).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 500 AA; 55446 MW; AE455E2780147DB8 CRC64;

Query Match 15.3%; Score 455; DB 1; Length 500;
Best Local Similarity 26.4%; Pred. No. 2.5e-24;
Matches 138; Conservative 85; Mismatches 225; Indels 74; Gaps 17;

QY 7 LSVPLVAASWA-----LP--GSTPASVGR-----OLPK--NPTGVKTLTANNVIRY 52
DQ 10 LLVVLAASACGLRLPRDAKFAQAERLIRSLNLLPKRAGTGGADVPVAPGELLE 69
QY 53 KPGAECVCTTPGVKSYG--VDTSPESHPTFFFEARNHNPETAPITLWNGPGSDS 110
DQ 70 RRVTLPGQGVGLGHAGYRLPNTHDARMEYFLFESRGKKED-PVWILWTGGPGCSS 128
QY 111 LIGLFEELGPGCHVNTFDDVINPHSMNEVSNLLFLSOPLGVSYSYSDTVGGINPTGVV 170
DQ 129 ELAVFYENGFTISNNMSLAWNKEGWDITINILFVDQPTGTGFSYSSDDRDTRHDETV- 187
QY 171 ENSFAGVQGRYPTIDATLIDTTNLAAEAAWEILQGLSLGSLDSRVQSKDFSLWTESY 230
DQ 188 -----SNDLYSFLQVFFKKHPEF-----AKNDPFITGESY 217
QY 231 GGHVPAFFNHFVEQNERIANGSVQVNFSLGIINGIIDEAIOAPYYPEFAVNNYTG 290
DQ 218 AGHYIFAFARVHGQGNK--AN--EGHINLKGFAIGNGLTDAIOVKAYTDVALD--- 268
QY 291 IKAVNETVVMYKMFANQMPGCGDILSTCKQNTALADYALCAEATNMCRDNVEGPIYA 350
DQ 269 MNLIKKSDYDRI---NKFIPPCFAFKLCTNGK-----ASCWAAVMVNCISFSSIMKL 319
QY 351 FAGRGYVDIRHPYDDPTPPSYN--KFLAKDSYMDAIGV--NINVTQSNNDVYVAFQQTGD 407
DQ 320 VGTNTYDVRKECEGKLCYDFSNLEKFFGDKAVKEALGVGDLEFVSCSTTVYQAMLTD-- 377
QY 408 FVWPNFIE-DLEILALPVRSLIYGDADYICNWFQGVQVSLAANYSOAAQFSAGYTPL 466
DQ 378 --WMRLVGVIPALLEDGINVLIVAGEYDLICNWLGNRSRWMSMWSGQKDFVSSHESFP 435
QY 467 KNGVEYGETREYGNFSFTRVVEAGHEVPVYQPIASLQLFNR 508
DQ 436 VVDGAEGVYKSHGSLFLKVNAGHVMVDQPKASLEMLR 477

RESULT 10
CBP3 HORVU STANDARD; PRT; 508 AA.
AC CBP3 HORVU
AC P21529;
DT 01-MAY-1991 (Rel. 18, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase III precursor (EC 3.4.16.5) (CP-MIII).
GN CBP3 OR CXP;3.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Himalaya; TISSUE=Aleurone;
RA Rocher A., Lok F., Cameron-Mills V., von Wettstein D.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 81-491.
RC STRAIN=cv. Gula;

```

```

RX MEDLINE=903115015; PubMed=2639682;
RA Soerensen S.B., Svendsen I., Breddam K.;
RL "Primary structure of carboxypeptidase III from malted barley.";
RL Carlsberg Res. Commun. 54:193-202(1989).
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
CC broad specificity.
CC -!- ENZYME REGULATION: INHIBITED BY MERCURIC IONS.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED MAINLY IN THE ALEURONE AND, TO A
CC LESSER EXTENT IN THE EMBRYO, THROUGHOUT THE 5-DAYS GERMINATION
CC PERIOD EXCLUSIVELY WITH A MAXIMAL LEVEL AT 3 DAYS. ALSO FOUND IN
CC THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y09604; CAA70817.1; .
CC DR HSP; P00729; ICFY.
CC DR MEROPS; S10.009; .
CC DR InterPro; IPR000379; Ser_estr site.
CC DR Pfam; PF00450; serine carboxpept; 1.
CC DR PRINTS; PR00724; CRBOXYPTASEC.
CC DR ProDom; PD001189; Serine carboxpept; 1.
CC DR PROSITE; PS00131; CARBOXYPEPT SER_SER; 1.
CC DR PROSITE; PS00560; CARBOXYPEPT SER_HIS; 1.
CC DR Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 80
FT CHAIN 81 491 SERINE CARBOXYPEPTIDASE III.
FT PROPEP 492 508
FT MOD_RES 81 81 BLOCKED.
FT ACT SITE 223 223 BY SIMILARITY.
FT ACT SITE 411 411 BY SIMILARITY.
FT ACT SITE 468 468 BY SIMILARITY.
FT BINDING 414 414 SUBSTRATE.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .).
FT VARIANT 265 265 Q -> V (POLYMORPHISM).
SQ SEQUENCE 508 AA; 56362 MW; 70C6751D78D40AB6 CRC64;

Query Match 14.7%; Score 437; DB 1; Length 508;
Best Local Similarity 26.5%; Pred. No. 4.5e-23;
Matches 134; Conservative 78; Mismatches 225; Indels 68; Gaps 15;

QY 12 LVAASWALFGSTPASVGRRLPKNTGVKTLTANNVTIRYKEPG-AGEVCETTGKSY 70
DQ 40 LIRALNLLPKDSSSSSGRHG-ARVCEGNEDEVAPGQLLERRVTLPLGPGVAD----LGH 94
QY 71 SCY--VDTSPESHPTFFFEARNHNPETAPITLWNGPGSDSLIGLFEELGPGCHVNTFD 128
DQ 95 AGYRLPNTHDARMEYFFFEESRGKKED-PVWILWTGGPGCSSLAVFVENGPFITANNMS 153
QY 129 DYINPHSMNEVSNLLFLSOPLGVSYSYSDTVGGINPTGVVNSFAGVQGRYPTIDAT 188
DQ 154 LVWNKFGWKDINSIIFVDQPTGTGFSYSSDDRDTRHDETV----- 194
QY 189 LIDTTNLAAEAAWEILQGLSLGSLDSRVQSKDFSLWTESYGGHYGAFNHFVEQNER 248
DQ 195 -----SNDLYDYLQVFFKKHPEF-----IKNDFFITGESYAGHYIPAFASRVHGQGNK 242
QY 249 IANGSVNGVQVNFSLGIINGIIDEAIOAPYYPEFAVNNYTGIKAVNETVVMYKMFANQ 308
DQ 243 -----NEGTHINLKGFAIGNGLTDAIOVKAYTDVALEMLNLIQADYERI-----NKF 290
QY 309 PNGCQDLISTCKQNTNR-TALADYALCAEATNMCRDNVEGPIYAFAGRGVYDIRHEYDPT 367

```

Query Match 14.0%; Score 416.5; DB 1; Length 500;
Best Local Similarity 24.6%; Pred. No. 1.2e-21;
Matches 131; Conservative 85; Mismatches 231; Indels 85; Gaps 16;

QY 7 LSVLPLVAAGSWALPGSTPASVGRRLPKNPTG-----VKLTLTANNVTIRKPKGAE 58
Db 1 MATTPRLASLLALLALCAAAGALRLPDPASFGQAQERLIRALNLLPGRPRGLGAGAE 60
QY 59 GVCE-----TTPGVKS-----YSGY---VDTSPESHPTFFWFFSARHNPETAPITLM 101
Db 61 DVAPQLLERVTLPLGLPEGVGDIHGHEAGYRLNTHDARMFYFFFSRGKKED-PVVIW 119
QY 102 LGGPGSDSLGLPEELGPGCHVNSTEDDYINPHSWNEVSNLLFLSQPLGVGSYSDTVDG 161
Db 120 LTGPGGSSSELAVFYENGPTFIANNMSLVNWKFGMDKLSNIIFVDPATGTGFSYSSDDRD 179
QY 162 SINPVTGVNENSPAGVQGRYPTIDATLIDTNTLAARAAWEILOGLFSLGSLDSRVOSK 221
Db 180 TRHDEAGV-----SNDLYDFLQVFFKKHEPF-----VKN 208
QY 222 DFLWTSYGYGHPAPENHFEYEQNERIANGSVNGVQLNFSNLGIIINGIIIDEAIQAYYP 281
Db 209 DFFITGESYAGHYTPAFASRVHQGNK-----NEETHINLKGPAICNGLTDPALQYKAT 263
QY 282 EFVANNVYGIKAVNETVYNNMKEFANQMPGCGODLISTCKQTNR-TALADYALCAEATNMC 340
Db 264 DYALD-----MMLIQADYDRI---NKFTPPCEFAIKLCTGDKASCMAAYVCNSIFNSI 316
QY 341 RDNVEGPPYAFAGRGVYDIRHPYDDPTPPSYN--KFLAKOSVMDAIGV-NINYTQSNND 397
Db 317 MKLVGTKNY-----YDVRKECEKLCYDFSNLEKFFCDKAVROAIGVGDIEFVSCSTS 369
QY 398 VYVAFQQTGDFWENFIR-DLEIILALPVYSLIYGDADYICNWFPGGQAVSLAANYSQAA 456
Db 370 VYQAMLTD-----WMENLEVGIPALLEDGINVLIVYAGEYDLICNLWLNRSRWVHSMWESGOK 425
QY 457 QFRSAGYPLKUNGVGEYGETREYGNFSTRVVEAGHEVPYQFIASLOLQNR 508
Db 426 DPAKTAESFLVDDAQAGVLKSHGALSFLKVNAGHVMWMDQPKAALMLRR 477

RESULT 12
CPVL HUMAN STANDARD; PRT; 476 AA.
ID CPVL HUMAN Q9H3G5; Q9HBL7; Q9HAR7; Q9HBA1;
AC Q9H3G5; Q9HBL7; Q9HAR7; Q9HBA1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable serine carboxypeptidase CPVL precursor (EC 3.4.16.-)
DE (Carboxypeptidase, vitellogenic-like) (Vitellogenic carboxypeptidase-like protein) (VCP-like protein).
GN CPVL OR VLP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21295045; PubMed=11401439;
RA Mahoney J.A., Ntoliou B., Dasilva R.P., Gordon S., McKnight A.J.;
RT "Cloning and characterization of CPVL, a novel serine
carboxypeptidase, from human macrophages";
RL Genomics 72:243-251(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Cho J.-J., Baik H.-H.;
RT "Cloning of VCP-like protein expressed in human heart and placenta";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,

291 IPPCEFAIKLCTGNGKASCMAYMVCNTIFNSIMKLWTKNY-----YDVRKECEGKL 343
368 PPBYN--KFLAKOSVMDAIGV-NINYTQSNNDVYAFQQTGDFWENFIR-DLEIILAL 423
344 CYDFSLEKFGDKAVRAQALGVGDIIEFVSCSTS YQAVLTD-----WNNLEVGIPALLED 399
424 PVYSLYIGDADYICNWFPGGQAVSLAANYSQAAQFRSAGYTPLVKUNGVGEYGETREYGNFS 483
400 GINVLIVYAGEYDLICNLWLNRSRWVHSMWESGOKDFAKTAESSFLVDDAQAGVLKSHGALS 459
484 FTRYEAGHEVPYQFIASLOLQNR 508
460 FLKVNAGHVMWMDQPKAALMLRR 484

RESULT 11
CBP3 WHEAT STANDARD; PRT; 500 AA.
ID CBP3 WHEAT P11515;
AC P11515;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase III precursor (EC 3.4.16.5) (CP-WIII).
GN CBP3.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007602; PubMed=2820978;
RA Gaudin D.C., Barker R.F., Jarvis M.G.;
RT "A gibberellin responsive wheat gene has homology to yeast
carboxypeptidase Y";
RL J. Biol. Chem. 262:13726-13735(1987).
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SUBUNIT: Monomer (Probable).
CC -!- INDUCTION: By gibberellic acid (GA).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; J02817; AAA34273.1; -;
FIR; A29412; A29412.
DR HSP; P00729; ICPY.
DR MEROPS; S10.009; -;
DR InterPro; IPR000379; Ser_estrs_site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CBPOXYPTASEC.
DR PRODOM; PD001189; Serine_carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydroxylase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 21
FT PROPEP 22 73 BY SIMILARITY.
FT CHAIN 74 484 SERINE CARBOXYPEPTIDASE III.
FT PROPEP 485 500 BY SIMILARITY.
FT ACT SITE 216 216 BY SIMILARITY.
FT ACT SITE 404 404 BY SIMILARITY.
FT ACT SITE 461 461 BY SIMILARITY.
FT BINDING 407 407 SUBSTRATE (BY SIMILARITY).
FT CARBOHYD 144 144 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 500 AA; 55334 MW; B2ACE10EF8484CDA CRC64;

RA Nagahari K., Sugano S., Isogai T.;
 RT "HRI human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A.
 RF
 RC
 RX
 RA MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC
 CC -!- FUNCTION: May be involved in the digestion of phagocytosed
 CC particles in the lysosome, participation in an inflammatory
 CC protease cascade, and trimming of peptides for antigen
 CC presentation.
 CC
 CC -!- TISSUE SPECIFICITY: Expressed in macrophages but not in other
 CC leukocytes. Abundantly expressed in heart and kidney. Also
 CC expressed in spleen, leukocytes, and placenta.
 CC
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF106704; AAC37991.2; -
 DR EMBL: AF282617; AAG14348.1; -
 DR EMBL: AK075433; BAC11618.1; -
 DR EMBL: BC016838; AAL16838.1; -
 DR HSSP: PI0619; IIVI.
 DR MEROPS: S10.003; -.
 DR Genew: HGNC:14399; CPVL.
 DR InterPro: IPR000379; Ser estrs site.
 DR InterPro: IPR001563; Serine carbppt.
 DR Pfam: PF00450; serine carbppt; 1.
 DR PRINTS: PR00724; CRBOXYPASEC.
 DR ProDom: PD001189; Serine carbppt; 1.
 DR PROSITE: PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE: PS00560; CARBOXYPEPT_SER_HIS; FALSE_NEG.
 DR KX Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
 DR SIGNAL 1 22
 DR PROPEP 23 ?
 DR CHAIN ?
 DR FT PROBABLE SERINE CARBOXYPEPTIDASE CPVL.
 DR ACT_SITE 204 204
 DR FT BY SIMILARITY.
 DR ACT_SITE 388 388
 DR FT BY SIMILARITY.
 DR ACT_SITE 448 448
 DR FT BY SIMILARITY.
 DR CARBOHYD 91 81
 DR FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 DR CARBOHYD 132 132
 DR FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 DR CARBOHYD 307 307
 DR FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 DR CARBOHYD 346 346
 DR FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 DR CONFLICT 25 25
 DR FT R -> H (IN REF. 2)
 DR CONFLICT 284 284
 DR FT L -> F (IN REF. 3 AND 4).
 DR CONFLICT 287 287
 DR FT H -> L (IN REF. 3)
 DR CONFLICT 398 398
 DR FT H -> R (IN REF. 3 AND 4).

CC EMBL: U28730; AAA68259.1; --
 DR PIR; T16606;
 DR HSP; F10619; LIVY;
 DR MEROPS; S10.002; --
 DR WormPep; K10B2.2; CE02009.
 DR InterPro; IPR00379; Ser estrs site.
 DR InterPro; IPR001563; Serine carboxypept.
 DR Pfam; PF00450; serine carboxypept; 1.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine carboxypept; 2.
 DR PROSITE; PS00131; CARBOXYPEPT SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 KW Hypothetical protein; Hydrolase; Carboxypeptidase; Glycoprotein;
 KW Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 470
 FT ACT SITE 169 169
 FT ACT SITE 380 380
 FT ACT SITE 441 441
 FT ACT SITE 132 132
 FT CARBOHYD 316 316
 FT CARBOHYD 396 396
 FT CARBOHYD 470 470
 SQ SEQUENCE 470 AA; 53158 MW; CCC2DACB75EF30FC CRC64;
 Query Match 11.9%; Score 355.5; DB 1; Length 470;
 Best Local Similarity 25.2%; Pred. No. 1.9e-17;
 Matches 140; Conservative 69; Mismatches 194; Indels 153; Gaps 22;
 QY 5 EFLSVLPVLAASWALPGSTPASVGRRLPKNPTGVKTUTANNVIRKPEAGVCHT 64
 DB 2 KLSLFLFVSSYFCLAAPATDKVNDLP-----GLTF----- 35
 QY 65 PCVKSYSGVDTSPESHFFFEARHNPETAPITLWNGPGSDSLIGLPEELGFCVHN 124
 DB 36 PDFHYGYLRAWTDKYLHWLTESRAPDQPLVWLNWNGPGSCSLDLIELGPFHVK 95
 QY 125 STFDD-YINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGVVENSFAGVGRYP 183
 DB 96 DFGNSIYYEYAWNKEFANVLFLESFAGVGYSTNPNLTVSDDEVSLEHN----- 144
 QY 184 TIDATLIDTNLAABAEWILQGLFSLDSRVQSKDFSLWTSYGHYGPAPFNHY 243
 DB 145 -----YMALDLFLSKFPYKGR-----DFWITGESYAGVYPTL----- 178
 QY 244 EQNERIANGSVNGVQVNLNENSLGILGII-----DEALQAPYPEFAVNNTYGIKAVNETVY 299
 DB 179 --AVRLNDKKNFP--NFKGVALGNALPNPNNTWVPY-----YYHALVRDDLY 226
 QY 300 NYMKFANQMPNGQDLISTCKQTNRTALADYALCABATNCRDNVEGPPYAFAGRVYDI 359
 DB 227 N-----DIARNCCNNIGTCD-----IYSKFFDPN--CRDKV--INALDGTNELNM 268
 QY 360 RHPYDPTPPSYN-----KFLAKSVMDAIGV-----NINVT-----QSNNDVYV--- 400
 DB 269 YNLVD-----VCYNYPTNLKAPFIERQMRLAVGLPARKNAATVFLCACTNNTHVYLNK 324
 QY 401 -----AFQQTGDFVWPNF-----IDLEILALPVRVSLIYGDADYICN 439
 DB 325 ADVKRLSLHPLSPFLPAMEECSQVGKHYVTHFNVIFEFQTMIAAGIKILVYNGVDVTACN 384
 QY 440 WFGQGVASLAANYSSQAQ-----FRSAGYPLKNGVYGETREYNFSFTRVYVAGH 492
 DB 385 SIMNQFLTSLNLTVLGEQEKVNEAWHYSQTGTAVAGF---QTKFAGNVDFLTVRGSCH 441
 QY 493 EYVYQPIASLQ-LFN 507
 DB 442 FVPEDKPKESQOMIFN 457
 RESULT 14
 CF22_HORVU STANDARD; PRT; 436 AA.
 ID CF22_HORVU

P55748;
 01-NOV-1997 (Rel. 35, Created)
 01-NOV-1997 (Rel. 35, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 Serine carboxypeptidase II-2 precursor (EC 3.4.16.6) (CP-MII.2)
 (Fragment).
 CXP;2-2.
 Hordeum vulgare (Barley).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 Triticeae; Hordeum.
 NCBI_TaxID=4513;
 [1]
 SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 STRAIN=cv. Alexis; TISSUE=Grain;
 MEDLINE=94336715; PubMed=7520177;
 Dal Degán F., Rocher A., Cameron-Mills V., von Wettstein D.;
 "The expression of serine carboxypeptidases during maturation and
 germination of the barley grain";
 Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213(1994).
 -- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
 or lysine residue.
 -- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
 COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 -- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. LOW
 LEVELS IN THE DEVELOPING ALEURONE AND EMBRYO. ALSO FOUND IN THE
 ROOTS AND SHOOTS OF THE GROWING SEEDLING.
 -- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING
 ENZYME MATURATION (BY SIMILARITY).
 -- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

 EMBL; X78878; CAB59202.1; --
 HSP; P08819; 1WHT.
 InterPro; IPR00379; Ser estrs site.
 InterPro; IPR001563; Serine carboxypept.
 Pfam; PF00450; serine carboxypept; 1.
 PRINTS; PR00724; CRBOXYPTASEC.
 ProDom; PD001189; Serine carboxypept; 1.
 PROSITE; PS00131; CARBOXYPEPT SER; 1.
 PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Multigene family.
 FT NON_TER 1 256
 FT CHAIN <1 256
 FT PROPEP 257 270
 FT CHAIN 271 436
 FT ACT SITE 149 149
 FT ACT SITE 350 350
 FT ACT_SITE 403 403
 FT DISULFID 56 333
 FT DISULFID 217 229
 FT DISULFID 253 281
 FT CARBOHYD 107 107
 SQ SEQUENCE 436 AA; 48952 MW; E0F82D97E0C34DC9 CRC64;
 Query Match 11.9%; Score 353.5; DB 1; Length 436;
 Best Local Similarity 25.7%; Pred. No. 2.4e-17;
 Matches 123; Conservative 76; Mismatches 173; Indels 107; Gaps 23;
 QY 70 YSGYVDTSPE--SHTFFWFEEARHNPETAPITLWNGPGSDSL-IGLFEELGFCVHNST 126
 DB 16 YAGYVTVSEDRGALFWFEAAHDPASKPELLWNGPGSCSSIAFGVGEVGFPHVND 75
 QY 127 FDD-YINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGVVENSFAGVGRPTI 185
 DB 76 GKGVHNPYSNWVANILFLDPSFVGVSYSNT-----SADLSN-----GDERTA 121


```

QY 186 DATLIDTNNLAEEAAWEILOGLSLPSLDGRVSKDFSLWTSYGGHYGPAFFNHFYEQ 245
Db 122 KDSLVLFTK-----W--LERF-----PQYKEREFLTGESVAGHYVQLAQAKRH 165
QY 246 NERANGSVNGVQLNFNSLGIINGIIDEAIOAPYPEFVAVNNTYGIKAVNETVYNYMKFA 305
Db 166 HEATGDKSI-----NLKGYMVGNAITDDF--HDHYGIFQYMWTTGL--ISDQTYKLLNIF 216
QY 306 NQMPNGC--QDLISCTQNTALADYALCAEANTMCRDNVEGP--YAFAG-----353
Db 217 -----CDFSFVHTSQCDK--ILDIA--STEAAGNIDSYSIFTPCHSSSFASRNKVKR 267
QY 354 -RGVYDRHPYDDPTPS---YNNKFLAKDS--YMDAIG-----VNINVTQSNND 397
Db 268 LRSVGKMGEGYDPCETKXHSIVYFNLEHVQKALHVNVPVIGSKWETCSEVINTNWKDCERS 327
QY 398 VYAFQQTGDFVFNFTEDLEILALPVRSVLSLYGDADYICNWFEGQAVSLAANYSOAAQ 457
Db 328 VLHIY-----HEIQYGLRIWMFSGDITDAV-----IPVTSTRYSIDA 364
QY 458 FRSGVYTPLKV---NGVEYCEYREYGNFSGFTRVVEAGHEVPPYQYPIASLOLENRTIFG 512
Db 365 LKLPVTPWEAWYDDDDGEVGTQGYKGLNFVTVRGAGHEVPLHRPKQALTLIKSLFAG 423

RESULT 15
NF31_NAEFO STANDARD; PRT; 482 AA.
AC P42661;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
OS Virulence-related protein Nf314 (EC 3.4.16.-).
OC Naegleria fowleri.
OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX NCBI_TaxID=5763;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEE;
RX MEDLINE=92267659; PubMed=1587609;
RA Hu W.-N., Kopathik W., Band R.N.;
RT "Cloning and characterization of transcripts showing
RT virulence-related gene expression in Naegleria fowleri.";
RL Infect. Immun. 60:2418-2424(1992).
CC -!- FUNCTION: MAY BE REQUIRED BUT IS NOT SUFFICIENT FOR INCREASED
CC VIRULENCE.
CC -!- INDUCTION: BY GROWTH ON MAMMALIAN CELLS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M88397; AAA29384.1; ..
DR PIR; A43828; A43828.
DR HSSP; P10619; 11VY.
DR MEROPS; S10.UPW; ..
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine carboxypept.
DR Pfam; PF00450; serine carboxypept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 2.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase.
FT ACT_SITE 163 163 BY SIMILARITY.
FT ACT_SITE 399 399 BY SIMILARITY.
FT ACT_SITE 459 459 BY SIMILARITY.

```

```

SQ SEQUENCE 482 AA; 53848 MW; 0B83049C1A8A0908 CRC64;
Query Match 11.8%; Score 351; DB 1; Length 482;
Best Local Similarity 25.0%; Pred. No. 4e-17;
Matches 139; Conservative 57; Mismatches 165; Indels 196; Gaps 23;
QY 55 PGAEGVCETTPGVKSYSGYV--DTSPESHITFFWFPEARHNPETAPITLWNGGFGSDSLI 112
Db 20 PGLSG-----NIGKSYTGILLANATRGYLFYWFESMRNPSQDPLVMTWNTGGFGCSLIG 75
QY 113 GLFEELGFCNVNSTFDDYI-NPHSWNEVSNLLFLSOPLGVCFSYSDTVGDSINPVTGVVE 171
Db 76 GEASEHGLFLVNDAGATITRNPSWNRVSNLLIYEQVGVGFSYSNSTD-----124
QY 172 NSSFAGVQGYPTIDATLIDITNL-----AAEAAWEILOGLSLPSLDGRVSKDFSLWT 227
Db 125 -----DYQNLDVQAASDMNNALRDLFRFPQFGR-----ETVLIG 161
QY 228 ESYGGHYGP-AFPNHFYEQNERIANGSVNGVQLNFNSLGI--NGIID-----272
Db 162 ESYGGVVVPTTAYN-----IVEGNGKQQQPVVNLVGLVGVGVTDAEADSNISIPMM 213
QY 273 ---EAIQAPYYPPEFAVNNTYGIKAVNETVYNYMKFANQMGCCODLISTCKQNTRTALAD 329
Db 214 KYHSLISIKYEE-----GYKACQGDY-----ANQNLPAQCQKFLT-----249
QY 330 YALCAEATNMCRDNVEGPYYAF-----351
Db 250 -----DSSN-AMGNI-NPYVIYDSCPWLGINLQOKLKTTCQMTFQVLDPKTQOPVKIHL 302
QY 352 -----AGRVYDIRHPYDDPTPPSY-YNKFLAKDSVMDAIGVN-----388
Db 303 FQMYKHGWSKRVANERNFAPRFTDAPCVPNOSIAKYFRRLDVQOALGVRRKTADPNGW 362
QY 389 -----INVTQSNNDV--YVAFQQTGDFVMPNFIEDLEILALP-VRVSLIYGDADYICN 439
Db 363 NICTGIINTQYVSTILPFYA-----KLLPHIRILVYSGDTDMVNV 403
QY 440 WFGGQAVSLAANYSOAAQFRSAGYTPKLVNGVVEYGETREY---GNFSFTRVVEAGHEVP 495
Db 404 GLGTQAAIDKQLQOETSSWRKTWEFDS-ALGTVVGVGVIRKFEKSGKGLTITVNGAGHMVP 462
QY 496 YYQPIASLOLENRTIFG 512
Db 463 LVKPDSAFYMFKNFIDG 479

```

Search completed: November 21, 2003, 18:00:07
Job time : 18 secs

This Page Blank (uspto)

Db 436 HEVPMYQPEAALSMFQTIWISS 456

RESULT 3

T43236

carboxypeptidase C (EC 3.4.16.5) precursor [validated] - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: carboxypeptidase Y

C:Species: Schizosaccharomyces pombe

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 08-Sep-2000

C:Accession: T43236; T37997

R:Lee, B.R.; Takeuchi, M.; Kobayashi, Y.

R:Tabuchi, M.; Iwihara, O.; Ohtani, Y.; Ohuchi, N.; Sakurai, J.; Morita, T.; Iwihara, J.

J. Bacteriol. 179, 4179-4189, 1997

A:Title: Vacuolar protein sorting in fission yeast: cloning, biosynthesis, transport, a

A:Reference number: 551516; MUID:95308538; PMID:7789719

A:Accession: S51516

A:Molecule type: DNA

A:Residues: 1-460 <LEE>

A:Cross-references: EMBL:D16519; NID:5556466; PIDN:BAA03966.1; PID:9995456

A:Experimental source: strain NRIC 1199

A:Accession: S78013

A:Molecule type: protein

A:Residues: 52-62,90-99,367-381 <LEB>

A:Accession: S78014

A:Molecule type: mRNA

A:Residues: 18-460 <LES>

C:Genetics:

A:Gene: spcZ

A:Introns: 64/3; 95/1; 107/2; 138/3; 234/2; 246/1; 270/1; 343/3; 353/2; 384/3; 437/3

C:Superfamily: serine carboxypeptidase

C:Keywords: glycoprotein; hydrolase; serine carboxypeptidase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-51/Domain: propeptide #status predicted <PRE>

F:52-460/Product: serine-type carboxypeptidase #status experimental <MAT>

F:40,46,119/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:186/Active site: Ser #status predicted

Query Match 25.1%; Score 749; DB 1; Length 460;

Best Local Similarity 35.7%; Pred. No. 4.4e-44;

Matches 179; Conservative 72; Mismatches 188; Indels 62; Gaps 15;

QY 12 LVAASWALPGSTPASVGRQLKPNPVGKLTITANNVIRYKEPGAEGVCETTPGVKYS 71

DB 18 LVCAPVTVQAHPMHVLRRQ-DGNDT-----SSGNTTQDKYSP---KLCD--PDVKQYS 66

QY 72 GYVDSPEHSTFWFEARHNPTAPITLWNGPGSDSLIGLFEELGPGCHVNSTPDDYI 131

DB 67 GLDAAANDRYFWFPEKNDPKNDPLTLWNGPGCSLLGLWEELGPGCQNGS---A 122

QY 132 NPHSMNEVSNLLFLSPLGVGFSYSTVDGSIINFTGVVENSFAGVQRYPTIDATLID 191

DB 123 NPHSMHSSNMLFPDQPDGVGFSY-----GKQ-----TVS 152

QY 192 TTNLAEEAAWEILQGLSLGSLDSRVQSKDFSLWTSYGGHYGPAFFNHFYQNBRIAN 251

DB 153 TTEDAARAWTFLOAFVETFP-----QVSKLDVHYFGESYGGHYIPGFASHVDMNKKVQS 208

QY 252 GSVNGVQLNFSNGIINGIIDEAIOAPYPEFAVNNYTGKAVNETVYNNMVFANOMPNG 311

DB 209 GEEKGVVPLKSLGVNGFIDAVIQKSPKMTCHSTYPAVLSEE---ECDKMQQIYEND 265

QY 312 CQDLISTCKGTNRTALADYALCAEATNMCRDNVEGYPYAGRGVYDIRHPYDDPTPSY 371

DB 266 CKPAEEQCAESD-----EDSDCVNANQC--GQIEG-IYAQSGYSFYDIRQQGDD--TPHPF 317

QY 372 YNKFLAKDSYMDAIGVNNYNTQSNNDVYAFQOTGDFWPNFTIEDLEELALPVRSLLIY 431

DB 318 VDE-LNKASVIVKEVGARGHSMCSDSVGTAFQTGCA-RSYIPAVEKLLKEGIPVLIYV 375

QY 432 GDADYICNWFPGQAVSLAANYSOAAQFRSAGYTPKLVNGVEYGETRYGNFSTRVEAG 491

DB 376 GDADVICNWNGLDVADSLKWDGDAFSKTKLEAWKADKEVGQFRSADKLTFFVRVYEAG 435

QY 492 HEVPMYQPEAALSMFQTIWISS 512

Db 436 HEVPMYQPEAALSMFQTIWISS 456

RESULT 3

T43236

carboxypeptidase C (EC 3.4.16.5) precursor [validated] - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: carboxypeptidase Y

C:Species: Schizosaccharomyces pombe

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 08-Sep-2000

C:Accession: T43236; T37997

R:Lee, B.R.; Takeuchi, M.; Kobayashi, Y.

R:Tabuchi, M.; Iwihara, O.; Ohtani, Y.; Ohuchi, N.; Sakurai, J.; Morita, T.; Iwihara, J.

J. Bacteriol. 179, 4179-4189, 1997

A:Title: Vacuolar protein sorting in fission yeast: cloning, biosynthesis, transport, a

A:Reference number: 551516; MUID:95308538; PMID:7789719

A:Accession: S51516

A:Molecule type: DNA

A:Residues: 1-460 <LEE>

A:Cross-references: EMBL:D16519; NID:5556466; PIDN:BAA03966.1; PID:9995456

A:Experimental source: strain NRIC 1199

A:Accession: S78013

A:Molecule type: protein

A:Residues: 52-62,90-99,367-381 <LEB>

A:Accession: S78014

A:Molecule type: mRNA

A:Residues: 18-460 <LES>

C:Genetics:

A:Gene: spcZ

A:Introns: 64/3; 95/1; 107/2; 138/3; 234/2; 246/1; 270/1; 343/3; 353/2; 384/3; 437/3

C:Superfamily: serine carboxypeptidase

C:Keywords: glycoprotein; hydrolase; serine carboxypeptidase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-51/Domain: propeptide #status predicted <PRE>

F:52-460/Product: serine-type carboxypeptidase #status experimental <MAT>

F:40,46,119/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:186/Active site: Ser #status predicted

Query Match 21.9%; Score 652.5; DB 2; Length 1002;

Best Local Similarity 33.8%; Pred. No. 5.9e-37;

Matches 165; Conservative 66; Mismatches 174; Indels 83; Gaps 15;

QY 49 TIRYKEPGAEGVCETTPGVKYSYGVYDTSFESHSTFWFEARHNPTAPITLWNGPGS 108

DB 570 TIRVKDSKPESLGIDT--VKQYTGILVDVEDRHLFEWFFESRNDPNDPVLWNGPGC 627

QY 109 DSLIGLFEELGPGCHVN-STFDDYINPHSMNEVSNLLFLSPLGVGFSYSTVDGSIINFT 167

DB 628 SSLTGLFMBELGFPSSINIELTKPEYNPHSMNSNASVFLDQPIINTGFSNGD--DSVLDTVT 695

QY 168 GVVENSFAGVQRYPTIDATLIDTTNLAEEAAWEILQGLSLGSLDSRVQSKDFSLWT 227

DB 686 -----AKQVYAFNLFFAKFP---QYAHLDFFHAG 713

QY 228 ESYGGHYGPAFFNHFYQNER-----IANG-SVNGVOLNFSNGIINGIIDEALQAPYPE 282

DB 714 ESYAGHYIQQFAKEIMEHNGANFFVAGSYEMEKQVINKSVLIGNGLTDLPLVQYFYFGK 773

QY 283 FAVNNTYG-----IKAVNETVYNNMVFANOMPNGQDLISTCKGTNRTALADYAL 332

DB 774 MACESPYPIMSQEEDCRITGAYDT-----CAKLITGCTGFT-----PV 814

QY 333 CASATNMCRDNVEGYPYAFAGRGVYDIRHPYDDPTPSY-----YNKFLAKDSYMDAIGV 387

DB 815 CIGASLYCNAMIGP-FTKGLMIYDIRCECRQEHLCYPETGAIESYLNQEFVQEAIGV 873

QY 388 NINYQSNNDVYAFQOTGDFWPNFTIEDLEEL--ALPVRVSLIY-GDADYICNWFPGQ 444

Db 874 EYDKGNTNTEWIGFLPKGDMWRKTRFDDVTAILEAGLPV---LIYAGDADYICNYMGNE 930
QY 445 AVSLAANYSOAQRAGYTFPKVNGVEYGETREYGNFSTRVYVYAGHEVPYOPTASLQ 504
Db 931 AWTDALEWAGQREYEAELKWPSPNGKEAGKSGKFNFGYLRILYBAGHWVPFNOPEASLE 990
QY 505 LFNRTIFG 512
Db 991 MLNSWIDG 998
RESULT 4
S46008
Probable carboxypeptidase C (EC 3.4.16.5) YBR139w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein YBR1015; hypothetical protein YBR139w
C:Species: *Saccharomyces cerevisiae*
C:Date: 26-Aug-1994 #sequence_revision 05-Jan-1996 #text_change 19-Apr-2002
C:Accession: S46008; S46581
R:Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45995
A:Accession: S46008
A:Molecule type: DNA
A:Residues: 1-508 <BEC>
A:Cross-references: EMBL:Z36008; NID:9536435; PIDN:CAA85097.1; PID:9536436; GSPDB:GN0000
A:Experimental source: strain S288C
R:Becam, A.M.; Cullin, C.; Grzybowski, E.; Lacroite, F.; Nasr, F.; Ozier-Kalogeropoulos,
Yeast 10(Suppl.A), S1-S11, 1994
A:Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 complete
A:Reference number: S46569; MUID:94378717; PMID:8091856
A:Accession: S46581
A:Molecule type: DNA
A:Residues: 1-508 <BE2>
A:Cross-references: EMBL:X75891; NID:9496856; PIDN:CAA53497.1; PID:9496869
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0000343
A:Map position: 2R
A>Note: MIPS:YBR139w
C:Superfamily: serine carboxypeptidase
C:Keywords: hydrolase; serine carboxypeptidase; transmembrane protein
F:6-22/Domain: transmembrane #status Predicted <TM>
F:219,415,474/Active site: Ser, Asp, His #status predicted
Query Match 21.3%; Score 636; DB 1; Length 508;
Best Local Similarity 32.6%; Pred. No. 3.2e-36;
Matches 167; Conservative 70; Mismatches 197; Indels 78; Gaps 14;
QY 19 LPGSTPASVGRQL-----PKNPTGVKTLTTANNVTIRYKEPGAEGVCETTFGVKSYGYV 74
Db 43 LPQNTQTLKLDRLNHDPLFTFTFISSVDYSLRLRTVDPKSLGI----DTVKQWGYM 98
QY 75 DTPSESTFFFFFEARHNHPETAPITLWNGPGSDSLGLFEELGPGCHVNSTFDDYINPH 134
Db 99 DYKDSKHFFYFFFSRNDPANDPILLWNGPGGSSFTGLLFLFELGPGSSIGADMKPIHPY 158
QY 135 SWNEVSNLLFLSQPLGVGFYSYSDTVGSDINPVTGVVNSFAGVQGRYPTIDATLIDTIN 194
Db 159 SWNNASMIFFLEQLGVGFYSYD-----EKVSTK 188
QY 195 LAEAANEILLOGLFSLGSLDSRVQSDPSLWTSYSGHYGPAPFNHFYEQN-ERIANGS 253
Db 189 LAGDAVIFELFFPEAFPHL-----RSNDFHTAGESYAGHYIPQIAHEIIVKNRPER---- 239
QY 254 VNGVQLNFNSLGIINGLIIDEAIOAPYPEFAPV-NNTYGIKAVNETVYVMKPFANQMGNC 312
Db 240 ----TFNLTSVINGLITDPLIQADYEPMA CGGGHVPVLSSECKMSKAAGR----- 290
QY 313 QDLISTCKQTNRTALADYAL--CAEATNMCRDNVGPYAFAGVGVDIRHPYDDP---- 366
Db 291 -----CRRLNKLCKYASKSLPCIVATACDSALLEIYIN-TGLNVIDIRGPCEDNSTDG 343
QY 367 ---TPPSYNNKFLAKSDVMDAIGNI-NYQTSNNDVYVAFQQTGDFVMPNFIEDLEEILA 422

Db 344 MCVTGLRYVDQYNNFFVEQETLGSVDHNSGCDNDVFTGLFTGDSKPFQOYIAELLN 402
QY 423 LPVRVSLIYGDADYICNWFGGQAVSLAANYSOAQRFSAGYTP--LKVNGVEYGETREYG 480
Db 403 HNIPLVLIYAGDKDYICNWLGHNHWSNELEWINKRRYQRRMLRPWVSKETGEELGOVKNYG 462
QY 481 NFSFTRYVYAGHEVPYQYPIASQLFNRTIFG 512
Db 463 PFTFLRIYDAGHWVPYDQPEASLEMVNSWISG 494
RESULT 5
JC7666
serine-type carboxypeptidase homolog precursor - *Emmericella nidulans*
N:Alternate names: carboxypeptidase Y homolog
C:Species: *Emmericella nidulans*, *Aspergillus nidulans*
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: JC7666
R:Ohsumi, K.; Matsuda, Y.; Nakajima, H.; Kitamoto, K.
Bioosci. Biotechnol. Biochem. 65, 1175-1180, 2001
A:Title: Cloning and characterization of the CPYA gene encoding intracellular carboxypep
A:Reference number: JC7666; MUID:21333188; PMID:11440134
A:Accession: JC7666
A:Molecule type: mRNA
A:Residues: 1-552 <OHS>
A:Cross-references: DBJ:AB051820
C:Comment: This protein is an intracellular carboxypeptidase, which is localized in vacu
F:131-132/Region: cleavage site, by Arg-Ile #status predicted
C:Genetics:
A:Gene: CPYA
A:Introns: 165/1
F:1-15/Domain: signal sequence (or 1-18, or 1-19) #status predicted <SIG>
F:16-131/Domain: pro-sequence, vacuolar targeting signal (or 19-131, or 20-131) #status
F:131-132/Region: cleavage site, by Arg-Ile #status predicted
Query Match 20.2%; Score 601; DB 2; Length 552;
Best Local Similarity 31.6%; Pred. No. 9.3e-34;
Matches 154; Conservative 77; Mismatches 175; Indels 82; Gaps 19;
QY 45 ANNVTIRYKEPGAEGVCETTFGVKSYGVVDTSPE-SHTFFWFEEARHNHPETAPITLMLN 103
Db 127 AVDLRIKKTDPSSLGI---DDVKQYTGILDONENKHLFWFFESRNDKNDPVPVLMN 183
QY 104 GPGSDSLIQLFEELGPGCHVNSTFDDYINPHSWNEVSNLLFLSQPLGVGFYSYSDTVDGS 163
Db 184 GPGCSSLTGLFELGPGSDIDENIKPVYVYVAMNSASVIFLDQPVNVGYSYS----- 236
QY 164 NPVTGVVNSFAGVQGRYPTIDATLIDTTLAAEAANEILLOGLFSLGSLDSRVQSKDF 223
Db 237 -----GSTVSDTV-AAGKDVYVALLTLFFKQFP-----EYAEQDF 269
QY 224 SLMTESYGGHYGPAPFNHFYEQNERIANGSVNGVQLNFNSLGIINGLIIDEAIOAPYPEF 283
Db 270 HIAGESYAGHYIPVFTSEILLSHQR-----NINLKSVLINGLTDGTQVYEYRPM 320
QY 284 AVNNTYGIKAVNETVYVMKPFANQMGNPG---CQDLISTCKQTNRTALADYALCAEATNM 340
Db 321 ACSEGG-GYPAVLDE-----SSCRSDNALGRCSQMIESCYNSESA-----WVCVPASIVC 369
QY 341 RDNVGPGYPAAGVGVDIRHPYDDPT-----PPSYNNKFLAKSDVMDAIGNVNI-NYQTS 394
Db 370 NNALLAPYQR-TGQNVYDVGKCEDESNLCYKMGVYVSEYLNKPEVRAAAGVGVGVSDC 428
QY 395 NNDVYVAFQQTGDFVMP--NFIEDLEEILALPVRVSLIY-GDADYICNWFGGQAVSLAAN 451
Db 429 NFDINRNLFGDWMPKYPHELVPGLLE--QIPV---LIYAGDADYICNWLGHNHWSNELEWINKRRYQRRMLRPWVSKETGEELGOVKNYG 483
QY 452 YSOAQRFSAGYTPLVK-----NGVEYGETREYGNFSTRVYVYAGHEVPYQYPIASQL 506
Db 484 WPGHKEFAAAPMEDLKIVDNEHTGKIGQIKTHGNFTFMYLYGGHVMVMDQPEASLEFF 543
QY 507 NRTIFG-W 513

Mon Nov 24 13:41:13 2003

Db 544 NRMLGGEW 551

RESULT 6

JC1380
 C;Alternate names: carboxypeptidase Y
 C;Species: Candida albicans
 C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
 C;Accession: JC1380
 R;Mukhtar, M.; Logan, D.A.; Kaeufer, N.F.
 A;Title: The carboxypeptidase Y-encoding gene from Candida albicans and its transcript
 A;Reference number: JC1380; MUID:93051356; PMID:1427093
 A;Accession: JC1380
 A;Molecule type: DNA
 A;Residues: 1-542 <MUK>
 A;Cross-references: GB:M95182; NID:gl70828; PIDN:AAA34326.1; PID:gl70829
 A;Note: the authors translated the codon GAT for residue 42 as Asn, AAA for residue 373
 C;Genetics:
 A;Gene: CPV1
 C;Superfamily: serine carboxypeptidase
 C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; zymogen
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-127/Domain: propeptide #status predicted <PRO>
 F;128-542/Product: carboxypeptidase C #status predicted <MAT>
 F;213,291/Binding site: carboxylate (Asn) (covalent) #status predicted
 F;269,461,518/Active site: Ser, Asp, His (status predicted)
 F;316-330,347-356/Disulfide bonds: #status predicted

Query Match 20.0%; Score 596; DB 1; Length 542;
 Best Local Similarity 31.5%; Pred. No. 2e-33;
 Matches 169; Conservative 64; Mismatches 193; Indels 110; Gaps 20;

QY 19 LPGSTP--ASVGRRLPKNPTGVKLTANNVT--IRYKEPGRAGVCE----- 63
 Db 71 LDGLTPEIKNIWLMKFP-----NSTELNFKAPPKKIKITQQDFHVDQAV 121

QY 64 -----TP-----GVKSYGYVDTSPE--SHFFWFFFAHRNPETAPITLWNGGFG 107
 Db 122 PNHKRIKSTPKDLGIDTVKQSGVLDVVDDEKHFYFFESRNDPKNDPVLWNGGFG 181

QY 108 SDSLGLPEELGPHVNSTFDDYINPHSWNEVSNLLFTSQPLGVGFSDTVGDSINPVT 167
 Db 182 CSSLGLPEELGPHSSIDKNLKPYNPHSWNASVIFLDQPINVGYSY----- 230

QY 168 GVENSFPAGVQGRPTIDATLIDTNLAABAAWELLQGLSLDSRVSQKDFSLWT 227
 Db 231 -----SQSVNSTIAAGKDVYAFLOLFFKNPP-----EYANLDFHIAG 267

QY 228 ESYGHHYGPAPFNHFEQNERIANGSVNGVQLNFSGLII--NGIIDEAIQAPYYPEFAV 285
 Db 268 ESYAGHYIAPASEILITHEP-----NFLTSLVINGLTDPLVQVEYEPMAC 316

QY 286 NNTYGHKAVNETVYNYKFNANQPNQCGDLISCKGTNTALADYALCHEATNCRDNYE 345
 Db 317 GEGGSPVLEPEECDDGM--LNLPLR--CLSLIESCYE-----GSVMSCVPATIIYCNQGM 368

QY 346 GPYAFAGRVYDIRPYDDP-----TPPSYNNKFLAKDSVMDAIGVNI--YTQSNNDVYV 400
 Db 369 GPYQK--TGRNVYDIRTWCEGSSLCYQLEVIDYDYLNLPEVKXALGEVDEYQSCNFDIR 427

QY 401 AFOQTGDFVMP--NFIEDLEILALPVRVSLTY--GDADYICNWFQGVQAVSLAANYSQAA 456
 Db 428 NFMFAGDWNKPYQKNVLDLLEK--ELPV--LLVAGDKDFICNWLGNQAWTNLEWSGSK 482

QY 457 QPESAGVTPLVNGVEYGETREYGNFSFTVYAGHEVVPYQPIASIQLENRTIFG 512
 Db 483 GFTKAPVKSWKVGNKAAAGEVKNYKHFTFLRVFGGHHMVPYDQENALDMWNRWISG 538

RESULT 7
S61713

carboxypeptidase C (EC 3.4.16.5) precursor - yeast (Pichia pastoris)
 N;Alternate names: carboxypeptidase Y
 C;Species: Pichia pastoris
 C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 16-Jun-2000
 C;Accession: S61713
 R;Ohi, H.; Ohtani, W.; Okazaki, N.; Furuhashi, N.; Ohmura, T.
 A;Title: Cloning and characterization of the Pichia pastoris PRC1 gene encoding carboxy
 A;Reference number: S61713; MUID:96381245; PMID:8789258
 A;Accession: S61713
 A;Molecule type: DNA
 A;Residues: 1-523 <OHI>
 A;Cross-references: EMBL:X87987; NID:gl171615; PIDN:CAA61240.1; PID:gl171616
 C;Superfamily: serine carboxypeptidase
 C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; zymogen
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-106/Domain: propeptide #status predicted <PRO>
 F;107-522/Product: carboxypeptidase C #status predicted <MAT>
 F;193,271,484,487/Binding site: carboxylate (Asn) (covalent) #status predicted
 F;249,441,498/Active site: Ser, Asp, His #status predicted
 F;296-310,327-336/Disulfide bonds: #status predicted

Query Match 20.0%; Score 594.5; DB 1; Length 523;
 Best Local Similarity 33.1%; Pred. No. 2.4e-33;
 Matches 153; Conservative 59; Mismatches 169; Indels 81; Gaps 16;

QY 68 KSYSGYVD--SPESHTFWFFFAHRNPETAPITLWNGGPGSDSLIGLFEELGPHVNST 126
 Db 121 KQSYGLDVEADDKHFFYWFESRNDPQNDPIILWNGGPGCSLTGLFELGSSRINEN 180

QY 127 FDDYINHSWNEVSNLLFSLQPLGVGFSDTVGDSINPVTGVVENSFAGVQGRPTID 186
 Db 181 LKPIFNPSWNGNIAIYLDQPVNVGFSYS----- 210

QY 187 ATLIDTTLAAEAAWELLQGLSLDSRVSQKDFSLMTESYGHYGPAPFNHFEYQON 246
 Db 211 SSSVSNTVAGEDVYAFLOLFFOHPP-----EYQTNDFHIAGESYAGHYIPVFAEDILSQK 266

QY 247 ERIANGSVNGVQLNFSGLII--NGIIDEAIQAPYYPEFAVNTYGIKAVNETVYNYMKF 304
 Db 267 NR-----NFLTSLVINGLTDPLQYRYEPMACGEG--GAPSVLP----- 306

QY 305 ANQMN-----GCQDLISCKGTNTALADYALCAEATNCRDNEVEYQAFAGGVYD 358
 Db 307 ADECECNMLVTQDKLSLIQACVDSQA-----FTCAPAALYCNNAQMGEPYQR--TGKNVYD 360

QY 359 IRHPYDDPT-----PPSYNNKFLAKDSVMDAIGVNI--YTQSNNDVYAPQQTGDFWPNF 413
 Db 361 IRKEDGGSLCYKLEFDITLQKQFVQDALGAEDVTYESCNFEINRNLFFAGDWMKP--Y 419

QY 414 IEDLEIL--ALPVRVSLTY--GDADYICNWFQGVQAVSLAANYSQAAQERSAGVTPLVKNG 470
 Db 420 HEHVSLLNKGHPV--LLVAGDKDFICNWLGNRAWTDLVFWDDAGFEKAEVQDWLVNG 476

QY 471 VYGYSTREYGNFSFTVYAGHEVVPYQPIASIQLENRTIFG 512
 Db 477 RKAGEFKYNSNTLYLVYDAGEWAPYDQENSHMWNRWISG 518

RESULT 8
CPBVY

carboxypeptidase C (EC 3.4.16.5) precursor [validated] - yeast (Saccharomyces cerevisiae)
 N;Alternate names: carboxypeptidase Y; protein YMR297W
 C;Species: Saccharomyces cerevisiae
 C;Date: 15-Oct-1982 #sequence_revision 31-Mar-1993 #text_change 15-Sep-2000
 C;Accession: A26597; S47458; A94609; A00909
 R;Wallis, L.A.; Hunter, C.P.; Rothman, J.H.; Stevens, T.H.
 A;Title: Protein sorting in yeast: the localization determinant of yeast vacuolar carb
 A;Reference number: A26597; MUID:87131100; PMID:3028649
 A;Accession: A26597
 A;Molecule type: DNA
 A;Residues: 1-532 <VAL>

A;Cross-references: EMBL:M15482; NID:g172238; PIDN:AAA34902.1; PID:g172239
 R;Barrell, B.G.
 Submitted to the EMBL Data Library, August 1994
 A;Reference number: S47445
 A;Accession: S47458
 A;Molecule type: DNA
 A;Residues: 1-532 <BAR>
 A;Cross-references: EMBL:X80836; NID:g1289327; PIDN:CAA56806.1; PID:g530354; GSPDB:GN000
 R;Svendsen, I.; Martin, B.M.; Viswanatha, T.; Johansen, J.T.
 Carlsberg Res. Commun. 47, 15-27, 1982
 A;Title: Amino acid sequence of carboxypeptidase Y. II. Peptides from enzymatic cleavage
 A;Reference number: A90763
 A;Accession: A90763
 A;Molecule type: protein
 A;Residues: 112-223, 'X', 225; 228-239, 'X', 241-259, 'HG', 262-267, 'X', 269-388, 'E', 390-451, 'N',
 R;Svendsen, I.
 submitted to the Atlas, October 1982
 A;Reference number: A94609
 A;Contents: disulfide bond
 A;Accession: A94609
 A;Molecule type: protein
 A;Residues: 224-227 <SV2>
 A;Note: this is a revision to the sequence in reference A90763
 R;Martin, B.M.; Svendsen, I.; Viswanatha, T.; Johansen, J.T.
 Carlsberg Res. Commun. 47, 1-13, 1982
 A;Title: Amino acid sequence of carboxypeptidase Y. Peptides from cleavage with cyanogen
 A;Reference number: A90762
 A;Contents: annotation; experimental details
 R;Endrizzi, J.A.; Remington, S.J.
 submitted to the Brookhaven Protein Data Bank, March 1994
 A;Reference number: A52472; PDB:1YSC
 A;Contents: annotation; X-ray crystallography at 2.8 angstroms; disulfide bonds
 C;Genetics:
 A;Gene: SGD:PRC1; LBC1; MIPS:YMR297w
 A;Cross-references: SGD:S0004912; MIPS:YMR297w
 A;Map position: 13R
 C;Superfamily: serine carboxypeptidase
 C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; yeast vacuole; zymogen
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-111/Domain: propeptide #status predicted <PRO>
 F;112-532/Product: carboxypeptidase C #status experimental <MAT>
 F;124,138,279/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F;167-409,304,328-351,335-344,373-379/Disulfide bonds: #status experimental
 F;257/Active site: Ser #status experimental
 F;449,508/Active site: Asp, His #status predicted
 Query Match 17.7%; Score 527.5; DB 1; Length 532;
 Best Local Similarity 29.1%; Pred. No. 1.1e-28;
 Matches 149; Conservative 71; Mismatches 195; Indels 97; Gaps 19;
 QY 33 PKNPTGVKT-----LTTANNVIRY-----KEPGAEGVCBTRGVKSYGVVDTSP- 79
 DB 82 PRFPAIKTKDMDFFVKNDALENQLRVNKDKPKILGI---DENVTOYGLDVEDED 138
 QY 80 SHTFWFTEARINPETAITLMLNGPGSDSLIGLFEELGPGCHVNSTFDDYNPNSWNEV 139
 DB 139 KHFFWFTEESRDPAKDPVILMLNGPGGCSLTGLFFELGPGSSIGDLKPIGNPYSWNS 198
 QY 140 SNLLFSLQPLGVGSYSDTVDGSIINPVTGVVNSFAGVQGYPTIDATLITTNLAABA 199
 DB 199 ATVIFLDQPVNVGFSYS-----GSSG-----VSNVAAAGKD 229
 QY 200 AWEILQGLSLGSLDSRVSQKDFSLWTESYGHHYGPAPFNHFYCNRIANGSVNGVL 259
 DB 230 VYNFLELFDDQPEYVNGKQ---DFHIAGESYAGHYIPFASILSHKOR----- 276
 QY 260 NPNLSGII--NGIIDEAIQAPYYPEFA-----VNNYGIKAVNETVYNTMKFANQP 309
 DB 277 NFNLTSVLIGNGLTDLPTQYNYEPMACGEGPSVLPSECSAMEDSL----- 325
 QY 310 NGCQDLISTCKQTNTALADYALCAEATNMCRDNVEGYPYAFAGRVYDIRHPYDD----- 365

DB 326 ERCLGLIESCYDSQ-----SWMSCVPATIIYCNNAQLAPYOR-TGRNVYDIRKCEGNNLC 379
 QY 366 -PTPPSYNKFSLAKDSVMDAIGVNNINVTQS--NRDYYVAFQQTGDFVWPNFIEDLEEIIAL 423
 DB 380 YPTLQD-IDDYLNQDYVKEAVGAEDVHYESCNFDINRNLFGADWMPK-YHTAVTDLLNQ 437
 QY 424 PVRVSLIYGADYICNVWFGQAVSLAANYSQAAQFRS---AGYTPLKVNGVEGETREYG 480
 DB 438 DLPILVYAGDKDFICNWLGNKAWTDVLPWKYDEFAFASQKVRNWTASITDEVA-GEVRSYK 496
 QY 481 NFSFTRVYAGHEVPYVQPIASLQLENRTIFG 512
 DB 497 HFTYLVFNGGHMVPDVPENALSMVNEWIHG 528
 RESULT 9
 T03607
 probable carboxypeptidase C (EC 3.4.16.5) cbp31 - rice
 N;Alternate names: serine-type carboxypeptidase homolog
 C;Species: Oryza sativa (rice)
 C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 21-Jul-2000
 C;Accession: T03607
 R;Washio, K.; Ishikawa, K.
 Plant Physiol. 105, 1275-1280, 1994
 A;Title: Organ-specific and hormone-dependent expression of genes for serine carboxypeptidase
 A;Reference number: Z14975; MUID:950627118; PMID:7972496
 A;Accession: T03607
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-429 <WAS>
 A;Cross-references: EMBL:D17587; NID:g409581; PIDN:BAA04511.1; PID:g409582
 A;Experimental source: cv. Yukihihikari
 C;Genetics:
 A;Gene: cbp31
 C;Superfamily: serine carboxypeptidase
 C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase
 F;76,414,417/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F;148,336,393/Active site: Ser, Asp, His #status predicted
 Query Match 15.9%; Score 473; DB 1; Length 429;
 Best Local Similarity 28.6%; Pred. No. 4.5e-25;
 Matches 130; Conservative 77; Mismatches 178; Indels 70; Gaps 16;
 QY 70 YSGY--VDTSPSHTFPFHFEARHPN-ETAPTILMLNGPGSDSLIGLFEELGPGCHVNST 126
 DB 17 HAGYVRLPNTHDARLFYFFESKSGKGDPPVIVLTGPGCGSSELALFYENGPHFIADN 76
 QY 127 FDDYINPHSWNEVSNLLFLSQPLGVGSYSDTVDGSIINPVTGVVNSFAGVQGYPTID 186
 DB 77 NSLVWDFGWDGSESLIYVDQGTGTFYSNPRTRHDEAG-VSNDLYA----- 125
 QY 187 AFLIDTTNLAABAAWEILQGLSLGSLDSRVSQKDFSLWTESYGHHYGPAPFNHFYCN 246
 DB 126 -----FLQAPFTEHPNF-----AKNDFYITGESYAGHYIPAFASRYKGN 165
 QY 247 ERIANGSVNGVQLNFSNGIINGIIDEAIQAPYPEFAVNNYTGKAVNETVYNTMKFAN 306
 DB 166 K-----NSEGHINILKGFAGNGLTDPAIQVKAITYDSLD-----MGLTKSQFNRI--N 213
 QY 307 QWPNCGQDLISTCKQTNR-TALADYALCAEATNMCRDNVEGYPYAFAGRVYDIRHP--- 362
 DB 214 KIVPTCELAIKLCTGTSITCLGAYVVC-----NLICSSIE---TIIGKNKYDIRKPCVG 266
 QY 363 ---YDDPTPPSYNKFSLAKDSVMDAIGV-NINYTQSNNDVYVAFQGTGDFWPNFIE-DL 417
 DB 267 SLCYD-----LSNMEKFLQKSVRESLGVGDIOFVSCSTVIQAML-----LDWMNLEVG 318
 QY 418 EEILALPVRVSLIYGDADIICNVFGQAVSLAANYSQAAQFRSAGYTPKLVNGVEYGETR 477
 DB 319 PELLENDIKVLIYAGEVDLICNWLNSRWVNSWMSWGKEAFVSSSEEFPTVDGKEAGILK 378
 QY 478 EYGNFSFTRVYAGHEVPYVQPIASLQLENRTIFG 512

Db 379 SYGSLPLKVDHAGHVMVMDQPKVALEMLRWTSG 413

RESULT 10

S22530 carboxypeptidase C (EC 3.4.16.5) precursor - rice

N;Alternate names: carboxypeptidase III

C;Species: Oryza sativa (rice)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 24-Nov-1999

C;Accession: S22530

F;Washio, K.; Ishikawa, K.

Plant Mol. Biol. 19, 631-640, 1992

A;Title: Structure and expression during the germination of rice seeds of the gene for a

A;Reference number: S22530; MUID:92329723; PMID:162776

A;Accession: S22530

A;Molecule type: DNA

A;Residues: 1-500 <WAS>

A;Cross-references: EMBL:S40458

C;Genetics:

A;Introns: 100/2; 156/3; 196/3; 244/3; 295/1; 324/3; 411/1; 455/3

A;Superfamily: serine carboxypeptidase

C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; zymogen

F;1-21/Domain: signal sequence #status predicted <SIG>

F;122-73/Domain: propeptide #status predicted <PRO>

F;74-484/Prodomain: carboxypeptidase C #status predicted <MAT>

F;485-499/Domain: carboxyl-terminal propeptide #status predicted <CPRO>

F;144/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;216,404,461/Active site: Ser, Asp, His #status predicted

Query Match 15.3%; Score 455; DB 1; Length 500;

Best Local Similarity 26.4%; Pred. No. 9.8e-24;

Matches 138; Conservative 85; Mismatches 225; Indels 74; Gaps 17;

QY 7 LSVLPVLAASMA---LP--GSTPASVGR-----QLPK--NPTGVKTLTANNVTRY 52

Db 10 LVVVLAASCAEGRLPRDAKFAQAERLRSNLNLPKEAGTGAQVPSVAGELLE 69

QY 53 KPGAGVCGCTTGFVKSGY--VDTSPESHFTFFWFPEARHNPETAPITLWNGGPGSDS 110

Db 70 RRVTLPLGQGVGLGHAGYRLPNTHDARFYLFSRGKKED--PVVIMLTGGPGCSS 128

QY 111 LIGLEFELGPGCHVNSTFDYINPHSWNEVSNLLFSLQPLGVGFSYSDTVGSGINPTGVV 170

Db 129 ELAVFYENGPFPTISNNMSLAWNKFGWDITISNIFVDQPTGTGFSYSSDRDRTRHDETGV- 187

QY 171 ENSSFAGVQGRYPTIDATLIDITNLAAEAAWEILOGLSLPSLDSRVQSKDFSLWTESY 230

Db 188 -----SNDLYSFLQVFFKKHDEF-----AKNDFFITGESY 217

QY 231 GGHYGPAPFNHFYQNERIANGSVQVQLNFSGLINGIIDEAIQAPYYPPEFVANNVTVG 290

Db 218 AGHYIDAFASRVHOGNK--AN---EGHINLKGFAINGLITDPAIQKAYTDYALD--- 268

QY 291 IKAVNETVYNYMKFANQMGCGODLSTCKQTNRALADYALCAEATNCRDNVEGPPYA 350

Db 269 MNLIKKSDYDRI---NKFTPPCFALKCTGNK-----ASCMAYVMCNSIFSIMKL 319

QY 351 FAGRGVYDIRHPYDDTPPSYNN--KFLAKDSVMDAIGV--NINYTOSNNDVYVAFQOTGD 407

Db 320 VGTKNYDVRKCEGKCLYDFSNLEKFFGDKAVKEAIGVGDLEFVSCSTTVYQAMLTJ-- 377

QY 408 FVWPNFIE--DLEILALPVRSVLIYGDADYICNWFQGVQAVSLAANYSOAAQPSAGYTP 466

Db 378 --WRRNLEVGIPALLEDGINVLIVAGEYDLICNWLGNRSRVHMEWSGQKDFVSSHPF 435

QY 467 KUNGVGEFTREYCNFSFTVYAGHEVPYQPIASIQLPNR 508

Db 436 VVDGAGVLUKSHGFLSLKLVHAGHVMVMDQPKASLEMLR 477

RESULT 11

T48977

carboxypeptidase-like protein Fl4D17.80 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 24-Oct-2000

C;Accession: T48977

R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Rudd, S.;

submitted to the Protein Sequence Database, April 2000

A;Reference number: Z25008

A;Accession: T48977

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-510 <JOR>

A;Cross-references: EMBL:AL353992; GSPDB:CN00061; ATSP:Fl4D17.80

A;Experimental source: cultivar Columbia; BAC clone Fl4D17

C;Genetics:

A;Gene: ATSP:Fl4D17.80

A;Map position: 3

A;Introns: 115/2; 171/3; 211/3; 259/3; 310/1; 341/3; 428/1; 472/3

C;Superfamily: serine carboxypeptidase

F;231,421,478/Active site: Ser, Asp, His #status predicted

Query Match 14.7%; Score 439; DB 2; Length 510;

Best Local Similarity 26.6%; Pred. No. 1.3e-22;

Matches 123; Conservative 81; Mismatches 189; Indels 69; Gaps 14;

QY 64 TPGVKSY---SGY--VDTSPESHFTFFWFPEARHNPETAPITLWNGGPGSDSLIGLFEEL 118

Db 93 SPVQDFGHAGYKLPNSKARMEYFFESRTN--KADPVVIMLTGGPGCSSELALFYEN 151

QY 119 GPCVNSTFDYINPHSWNEVSNLLFSLQPLGVGFSYSDTVGSGINPTGVVENSFPAGV 178

Db 152 GPFTVSNSSLSWNEFGWMDKASNLTYVDQPTGTGFSYSDQSLRHDSDGV----- 202

QY 179 QGRYPTIDATLIDITNLAAEAAWEILOGLSLPSLDSRVQSKDFSLWTESYGGHYGPAF 238

Db 203 -----SNDLYDFOAFFKEHP-----QFVNDFYITGESYAGHYIPAL 240

QY 239 FNHFYQNERIANGSVQVQLNFSGLINGIIDEAIQAPYYPPEFVANNVTVGIKAVNETV 298

Db 241 ASRVHGNK-----NKEGTHINLKGFAINGLITDPAIQKAYADYALDNLITQSDHNL 295

QY 299 YNYMKFANQMGCGODLSTCKQTNRALADYALCAEATNCRDNVEGPPYAFAGR- 354

Db 296 NRY-----YATCQSIKECSADGEGDACCASYTVC--NNIFQKIMDIAGNV 340

QY 355 GYVDIRHPYDDTPPSYNN--KFLAKDSVMDAIGV--NINYTOSNNDVYVAFQOTGDVFWP 411

Db 341 NTYDVRKCEGSLCYDFSNMENFLNOKSVRKALGVGDIEFVSCSTAVIYFAMQND---- 396

QY 412 NFIE--DLEILALPVRSVLIYGDADYICNWFQGVQAVSLAANYSOAAQPSAGYTP 470

Db 397 RNLEVGIPALLODGIKLLVAGEYDLICNWLGNRSRVHMEWSGQKEFVAAATVPPHVDN 456

QY 471 VYGETREYCNFSFTVYAGHEVPYQPIASIQLPNRITFG 512

Db 457 KEAGLMKNYGSITFLKLVHAGHVMVMDQPKAALQMLQNMQ 498

RESULT 12

A35275

carboxypeptidase C (EC 3.4.16.5) - barley

N;Alternate names: carboxypeptidase III

C;Species: Hordeum vulgare (barley)

C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999

C;Accession: A35275

R;Sorensen, S.B.; Svendsen, I.; Breddam, K.

Carlsberg Res. Commun. 54, 193-202, 1989

A;Title: Primary structure of carboxypeptidase III from malted barley.

A;Reference number: A35275; MUID:90315015; PMID:2639682

A;Accession: A35275

A;Molecule type: protein

A;Residues: 1-411 <SCR>

C;Superfamily: serine carboxypeptidase

C;Keywords: blocked amino end; glycoprotein; hydrolase; serine carboxypeptidase

F:1/Modified site: blocked amino end (Leu) (probably acetylated) #status experimental
F:71/Binding site: carboxylate (Asn) (covalent) #status experimental
F:145,331,388/Active site: Ser, Asp, His #status predicted

Query Match	14.3%;	Score	425.5;	DB	1;	Length	411;
Best Local Similarity	26.5%;	Pred. No.	8.1e-22;				
Matches	121;	Conservative	71;	Mismatches	203;	Indels	62;
Gaps	12;						
Qy	59	GVCETTPGVKSYGY--VDTPESHTFPWFPEARHNPTAPITLWNGPGSDSLIGLIFE	116				
Db	3	GLPEGVADLGHAGYIRLPHDARMFYFFPESRGKED-PVWILWTGPGCSSLAVFY	61				
Qy	117	ELGCHVNSTFDDYNPHSMWENLLFLSQPLGVGHSYSDTVGSGINFTVGTGVNNSFA	176				
Db	62	ENGFTIANMNSLVNKGWDKISIIIVDQFTGTGFSYSSDDRDTRHDTGV-----	114				
Qy	177	GVQGRYPTIDATLDTTNLAEEAAWEILQGLSLDSRVQSKDSLWTESYGHGYP	236				
Db	115	-----SNDLYDFLOVFFKKHPEF-----	150				
Qy	237	AFNHFVEQNERIANGSVGQLNFNSLGIINGIIDEAIOAPYYPEFAVNNTYGIKAVNE	296				
Db	151	AFASRVHQGNKK-----NEGTHINLKGFAIGNGLTDPALYKAYTDVALEWNLLOKADYE	205				
Qy	297	TVVNYMKEAFQMPNGCCDLISCTQOTNR-TALADYALCABATNCRDNVEGYPYAFAGRG	355				
Db	206	RI-----NKFITPPCCFAIKLCTNGKASCMAAYVVCNTIFNSIMKLVTQNY-----	252				
Qy	356	VYDIRHPYDQTPPSYN--KFLAKDSVMDAIGV-NINYTCNSNDVYVAFQQTGDFWPN	412				
Db	253	-YDVKCEGKLCVDFSNLSEKFFGDKAVRQAIGVGDIIEFVSCSTSVTQAMLTD-----	307				
Qy	413	FIE-DIEEIIALPVRVSLIYGDADYICNWFPGQAVSLAANYSQAAQFSAGYTFLKNGV	471				
Db	308	NLEVGIPALLEDGINVLIIYAGEYDLICNLWLNSEWHSMWESGQOKFAKTAESSFLVDDA	367				
Qy	472	EYGETREYNFSTRVYVAGHEVYVYQPIASLQLFNR	508				
Db	368	QAGVLKSHGALSFLKVNHAGHWYPMQPKAALFMIPE	404				

RESULT 13

T18968
probable serine-type carboxypeptidase (EC 3.4.16.-) Y16B4A.2 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
C/Accession: T18968; T23145; T26477
R/Thomas, K.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19053
A/Accession: T18968
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2105 <WIL>
A/Cross-references: EMBL:Z70203; PIDN:CAA94110.1; GSPDB:GN000028; CESP:Y16B4A.2
R/Illoyd, C.
submitted to the EMBL Data Library, March 1997
A/Reference number: Z19697
A/Accession: T23145
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2105 <W12>
A/Cross-references: EMBL:Z93339; PIDN:CAB07544.1; GSPDB:GN000028; CESP:Y16B4A.2
R/Wallis, J.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z20220
A/Accession: T26477
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2105 <W13>
A/Cross-references: EMBL:AL023825; PIDN:CAA19443.1; GSPDB:GN000028; CESP:Y16B4A.2

A;Experimental source: clone Y16B4A
C;Comment: This protein has 4 repeats of sequence homologous to serine carboxypeptidase
C;Genetics:
A;Gene: CESP:Y16B4A.2
A;Map position: X
A;Introns: 43/3; 67/2; 179/1; 208/3; 239/2; 281/2; 296/3; 346/2; 444/3; 503/3; 588/3; 616/4/3; 1752/2; 1794/3; 1834/3; 1882/2; 1905/2; 1942/3; 1982/3
C;Keywords: duplication; hydrolase; serine carboxypeptidase

Query Match	14.0%;	Score	417.5;	DB 2;	Length	2105;	
Best Local Similarity	23.5%;	Pred.	No. 2.8e-20;				
Matches	156;	Conservative	91;	Mismatches	209;	Gaps	28;

Qy	2	RGYEFLSVLPVLAASWA-----LPGSTPASVGRRRLPKNPITGVKLTITANN	47
Db	1467	QGSOYVTIIAGYAKSWTQLNVLTKVGSGHFVPSDRPAQ-----ALQMILTFLS	1515
Qy	48	VTRIYKEPGAEGV-----CET-----TPGV-----KSYSGYVDTSPE	80
Db	1516	NQANYSTPAGIDVTQPQLTNWISQTNCTTGQTRINLPGPLMDLFQKYSGLDLSGH	1575
Qy	81	HTEFWFEARHNPETAPITILWLNGGPGSDSLIGLFEELGFCHEVNSTEDDYI-NPHSMNEV	1339
Db	1576	KVHYWLVESENPTDPLLWLNLGGPGSSLMGLFEENGFFRVSKDSQTLRSRPYSWNKF	1635
Qy	140	SLLFLSQPLGVGFYSYSDTDVGSINPVTVGVENSFAGVOGRYPITDATIIDTTNLAAEA	199
Db	1636	ANVLYESPIGVGVSVA-----YNNINIQQDDVT--TAQE	1668
Qy	200	AWEILOQFSLGPLSDSRVOSKDFSLWTESYGCHYGPAFNHFHVEQNERLANGSVNG-VQ	258
Db	1669	NYAALKSFQAQYP----QYTTSFYTTGESYAGVYLPLGL-----SALLVQGIKSGDIN	1717
Qy	259	LNFNSLGINGIIDE-----AIQAPYYEPFAVNNTYGIKAVNETVYN-----	300
Db	1718	INYKGVSIGNGVIDKRKTDMNSQLHYQYY-----HG-GISATTYNTALQLCCSGDEF	1767
Qy	301	-----YMKFANQMENG-----CQD-LISTCKQNRATADYALCAEATNMCRDNVEGP	347
Db	1768	KCFSDRMTNFNNISIPWGDLSPCYDFVVATGAQLLLTAFDPVMYQOCWTI-----	1819
Qy	348	YVAFAGRGVVIDRHYPDDTPBPSYNNK-----FLAKOSV	381
Db	1820	-----PYNDTPRTPYGETWTGINYESDDLNGPCYDDSAMGEYLNRPPV	1865
Qy	382	MDAIGV--NINTQSNNDDVYAFQQTGDPFWFNFIEDLEEILA---LPVRVSLVIGDADY	436
Db	1866	RXALNIIPDSVPVYAANNIINAYNQVDSIVN---LQIIMANAPANFKMLLYSGDAIT	1921
Qy	437	ICNWFGGQ-----AVSLAANYGQAA---QFRSAGYTPLVKNVGEYGETREYGNFSFT	485
Db	1922	MVNLUGAEIFTANNFAALGLTSSPSRAOWTFQYNST-FQP-TVAGYQTSYTSNAINIDLVI	1979
Qy	486	RYVEAGHEYPPYOPTIASIQLEFNRTIFGWDIAGQKKIWPSKYKNCATATATHQTSSVP--L	543
Db	1980	TKVSGSHFPDLRPOCALQM-----LYNFVKRGYNTPFDLNSNFTHTTTTTTPGTG	2035
Qy	544	PTATS	548
b	2036	PTVTA	2040

RESULT 14

carboxypeptidase C (EC 3.4.16.5) precursor - wheat
N/Alternate names: carboxypeptidase Y homolog
C/Species: Triticum aestivum (common wheat)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999
C/Accession: A29412
R/Baulcombe, D.C.; Barker, R.F.; Jarvis, M.G.
J. Biol. Chem. 262, 13726-13735, 1987
A/Title: A gibberellin responsive wheat gene has homology to the carboxypeptidase Y gene
A/Reference number: A29412; MUID:98007602; PMID:28407114

A;Accession: A29412
A;Molecule type: mRNA
A;Residues: 1-499 <BAU>
C;Genetics:
A;Introns: 100/2; 156/3; 196/3; 244/3; 295/1; 324/3; 411/1; 455/3
C;Superfamily: serine carboxypeptidase
C;Keywords: glycoprotein; hydrolyase; serine carboxypeptidase; zymogen
F;1-22/DNA: signal sequence #status predicted <SIG>
F;23-73/DNA: signal sequence #status predicted <SIG>
F;74-484/Product: carboxypeptidase C #status predicted <MAT>
F;485-499/DNA: carboxyl-terminal propeptide #status predicted <CPRO>
F;144/Binding site: carboxyl-terminal (Asn) (covalent) #status predicted
F;216,404,461/Active site: Ser, Asp, His #status predicted

Query Match 12.9%; Score 383; DB 2; Length 482;
Best Local Similarity 26.2%; Pred. No. 8.7e-19;
Matches 135; Conservative 68; Mismatches 202; Indels 110; Gaps 18;

QY 33 PKNPTGVKLTANNVIRYKE-----PCAEGVCETTPG-----VKSYSYGVVDTSPEST 82
DB 32 PPKRGVSSGDTSHENVIRQESVPKDKLIQOLPGQSDVTFKQGGVAVNKKPGR 91
QY 83 F--FWPEAREARNPTAPITLLMNGPGSDSLIGLFEELGPGCHVNSTFDD--YINPHSWNEV 139
DB 92 FLYIFVETIKPGNTPLVIMFNGGPGCSLGGAFKELGPRFVHSDGKTLFRNPYSWNE 151
QY 140 SNLLFLSQPLGVGSYSYDVGSIINVTGVVENSFAGVQGRYPTIDATLIDTINLAEEA 199
DB 152 ANVLFLFETPVGTFGSYS-----NSPINGKQDKAT-----ABD 184
QY 200 AWEILQGLSLGLPSLDSRVQSKDFSLWTSYGGHYGPAFFNHFYEQNERIANGSVNGVQL 259
DB 185 NYMFLVNLERFPPEYKGR---DIVIAGQSVAGHYVVPQLAQILHFN-----NQTLI 232
QY 260 NENSLGILNGIIDEAIAQAPYYPEPAVNNTYGIKAVNETVYNYMKEANQMPNGCODLISCT 319
DB 233 NLRGLIGNPSLNREIQDDFGYKMF--SHGLIS--QQQMDNYNKF-----CTD--SDL 280
QY 320 KOTNRATADYALCAEATNMCRDNVEGPPYAFAGRGVDIRHPY-----DIYNIYAPLCINSLSEPKKCTTI 325
DB 281 YDWDKCHLASQKIEAQKTHL-----INVTQSNNDVYVAFQQTGDFWPNFIED 416
QY 364 --DDTPPSYNNKFLAKDSVMDAIGVN-----INVTQSNNDVYVAFQQTGDFWPNFIED 416
DB 326 MKADPCSGNYLKAYLNKEVQBAIHANTTKIPEWTSCTNKLWENNEKDRIV--SLTFI 383
QY 417 LBEILALPVRVSLIYGDADYICNWFQGOAVSLAANYSOAQRSAGYTPKLVNGVEYGET 476
DB 384 LOELMGKGVVRLYNGVDVLPFTSTLAVVKTMLTVVKEWR-----PWFTHGHVGGFT 438
QY 477 REY-GNFSFTRVYEAGHEVPPYQPIASLQLFNRTI 510
DB 439 EDYKGNLTFTVTKGAGHSVPTDQPIHALNIFTSFI 473

Search completed: November 21, 2003, 18:01:33
Job time : 22 secs

A;Accession: A29412
A;Molecule type: mRNA
A;Residues: 1-499 <BAU>
C;Genetics:
A;Introns: 100/2; 156/3; 196/3; 244/3; 295/1; 324/3; 411/1; 455/3
C;Superfamily: serine carboxypeptidase
C;Keywords: glycoprotein; hydrolyase; serine carboxypeptidase; zymogen
F;1-22/DNA: signal sequence #status predicted <SIG>
F;23-73/DNA: signal sequence #status predicted <SIG>
F;74-484/Product: carboxypeptidase C #status predicted <MAT>
F;485-499/DNA: carboxyl-terminal propeptide #status predicted <CPRO>
F;144/Binding site: carboxyl-terminal (Asn) (covalent) #status predicted
F;216,404,461/Active site: Ser, Asp, His #status predicted

Query Match 14.0%; Score 416.5; DB 1; Length 499;
Best Local Similarity 24.6%; Pred. No. 4.4e-21;
Matches 131; Conservative 85; Mismatches 231; Indels 85; Gaps 16;

QY 7 LSVLPLVAASWALPGSTPASVGRROLKPNPTG-----VKTLTANNVIRYKEPGA 58
DB 1 MATTPRLASULLLLALCAAAGALRLPPDASPPGAERLIRALNLLPFRPRGLGAGAE 60
QY 59 GVCE-----TTPGVKS-----YSCY--VDTSPESHTFFWFEAREHNPTAPITLW 101
DB 61 DVAPQQLLERRVTLFCLPEGVDLGHAGYXPLNTHDARMEYFFESRGKKED--PVVILW 119
QY 102 LMGPGSDSLIGLFEELGPGCHVNSTFDDYINPHSWNEVSNLLFLSQPLGVGSYSYDVG 161
DB 120 LTGGPGCSSELAFFVENGFTIANNSLVNWKFGWKISIIIFVDPATGTGFSYSSDDR 179
QY 162 SINPTGVVENSFAGVQGRYPTIDATLIDTINLAEEAWEILQGLSLGLPSLDSRVQSK 221
DB 180 TRHDEAGV-----SNLDYDFLQVFFKKHPEF-----VKN 208
QY 222 DFLSLWTSYGGHYGPAFFNHFYEQNERIANGSVNGVQLNPSLGIINGIIDEAIAQAPYP 281
DB 209 DFFITGESYAGHYIPAFASRVHOGNKK-----NEGTHINLKGFALNGLTDPAIQKAT 263
QY 282 EFANVTYGIKAVNETVYNYMKEANQMPNGCODLISCTKQTR--TALADYALCAEATNMC 340
DB 264 DYALD-----NMLIQADYDRI---NKFIPPCPAIKLCGTGDKASCWAAVWVCNSIFNSI 316
QY 341 RNVEGPPYAFAGRGVDIRHPVDDTPPSYN--KFLAKDSVMDAIGV-NINYTOSNND 397
DB 317 MKLVGTKNY-----YDVRKECEKGLCYDFSNLKFFGDKAVRQAIGVGDIEFVSCSTS 369
QY 398 VYVAFQQTGDFWPNFIE--DLEILALPVRVSLIYGDADYICNWFQGOAVSLAANYSOAA 456
DB 370 VYQAMLTLD---WMRLNLEVGIPALLEDGINVLITYAGEYDLICNLWGNRSRWHWSHWSGQK 425
QY 457 QRSAGVTPKLVNGVEYGETREYGNFSFTRVYEAGHEVPPYQPIASLQLFNRTI 508
DB 426 DPAKTAESSFLVDDAQAGVLKSHGALSFLKVHNGHVPMDQPKAALEMLRR 477

RESULT 15
T49079
serine-type carboxypeptidase like protein - Arabidopsis thaliana
N;Alternate names: protein F4P15.110
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
R;Alcaraz, J.P.; Clabault, G.; Cortet, A.; Mache, R.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000
A;Accession: T49079
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-482 <ALC>
A;Cross-references: EMBL:AL049711; GSPDB:GN00061; ATSP:F4P15.110
A;Experimental source: cultivar Columbia; BAC clone F4P15
C;Genetics:
A;Gene: ATSP:F4P15.110

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: November 21, 2003, 17:59:41 ; Search time 36 Seconds
 (without alignments)
 2814.459 Million cell updates/sec
 Title: US-09-712-338-2
 Perfect score: 2979
 Sequence: 1 MRGVEFLSVLPLVAASWALP.....HTQSSVPLTATSMSSVGM 555
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 666188 seqs, 182559486 residues
 Total number of hits satisfying chosen parameters: 666188
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries
 Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	517.5	17.4	421	9	US-09-420-785A-4
2	517.5	17.4	421	9	US-09-901-252-15
3	361	12.1	476	14	US-10-084-018-3
4	359	12.1	476	10	US-09-909-320-164
5	359	12.1	476	10	US-09-909-088B-164
6	359	12.1	476	10	US-09-905-291A-164
7	359	12.1	476	10	US-09-902-853-164
8	359	12.1	476	10	US-09-907-824-164
9	359	12.1	476	10	US-09-907-841-164
10	359	12.1	476	11	US-09-904-011-164
11	359	12.1	476	11	US-09-906-742-164
12	359	12.1	476	11	US-09-906-838-164
13	359	12.1	476	11	US-09-907-613-164
14	359	12.1	476	11	US-09-907-942-164
15	359	12.1	476	11	US-09-796-753-40

16	359	12.1	476	11	US-09-904-859-164	Sequence 164, App
17	359	12.1	476	11	US-09-909-204-164	Sequence 164, App
18	359	12.1	476	11	US-09-904-820-164	Sequence 164, App
19	359	12.1	476	11	US-09-904-786-164	Sequence 164, App
20	359	12.1	476	11	US-09-906-646-164	Sequence 164, App
21	359	12.1	476	11	US-09-906-700-164	Sequence 164, App
22	359	12.1	476	11	US-09-903-786-164	Sequence 164, App
23	359	12.1	476	11	US-09-903-749A-164	Sequence 164, App
24	359	12.1	476	11	US-09-902-903-164	Sequence 164, App
25	359	12.1	476	11	US-09-904-119-164	Sequence 164, App
26	359	12.1	476	11	US-09-904-956-164	Sequence 164, App
27	359	12.1	476	11	US-09-902-736-164	Sequence 164, App
28	359	12.1	476	11	US-09-907-794-164	Sequence 164, App
29	359	12.1	476	11	US-09-903-943-164	Sequence 164, App
30	359	12.1	476	11	US-09-904-462-164	Sequence 164, App
31	359	12.1	476	11	US-09-907-925-164	Sequence 164, App
32	359	12.1	476	11	US-09-902-692-164	Sequence 164, App
33	359	12.1	476	11	US-09-903-520-164	Sequence 164, App
34	359	12.1	476	11	US-09-905-056-164	Sequence 164, App
35	359	12.1	476	11	US-09-909-064-164	Sequence 164, App
36	359	12.1	476	11	US-09-904-553-164	Sequence 164, App
37	359	12.1	476	11	US-09-905-381-164	Sequence 164, App
38	359	12.1	476	11	US-09-905-088-164	Sequence 164, App
39	359	12.1	476	11	US-09-907-573-164	Sequence 164, App
40	359	12.1	476	11	US-09-905-075-164	Sequence 164, App
41	359	12.1	476	11	US-09-902-634-164	Sequence 164, App
42	359	12.1	476	11	US-09-902-713-164	Sequence 164, App
43	359	12.1	476	11	US-09-907-979-164	Sequence 164, App
44	359	12.1	476	11	US-09-902-615-164	Sequence 164, App
45	359	12.1	476	11	US-09-902-615-164	Sequence 164, App

ALIGNMENTS

RESULT 1
 US-09-420-785A-4
 ; Sequence 4, Application US/09420785A
 ; Patent No. US20010010923A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MORTENSEN, UFFE
 ; APPLICANT: OLESEN, KJELD
 ; APPLICANT: STERNICK, HENNING
 ; APPLICANT: SORESENSEN, STEEN B.
 ; APPLICANT: BREDDAM, KLAUS
 ; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
 ; FILE REFERENCE: 089187/0109
 ; CURRENT APPLICATION NUMBER: US/09/420,785A
 ; CURRENT FILING DATE: 1999-10-19
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 421
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-09-420-785A-4

Query Match	17.4%	Score 517.5	DB 9	Length 421
Best Local Similarity	29.5%	Pred. No. 6e-40		
Matches 142	Conservative 70	Mismatches 185	Indels 85	Gaps 17
QY	51	RYKFGAGVCETTPGVKSYGYVDTSP-SHTFFWFPEARNHNPETAPITLWNGRQSD	109	
Db	1	KIKDPKILGI---DENVVTQYGLDVEDEKHFHFWTFESRNDPAKDPVILWNGPGCS	57	
QY	110	SLIGLFEELGCHVNSTFDDVINPHSNVNSNLFSLQPLGVGFYSYTPDGSINPVTGV	169	
Db	58	SLTGLFELGESSIGDPLKPIGNEFYSNWSNATVFLDQPVNVVGFYS	104	
QY	170	VENSSFAGVCGRYPTIDATLIDTTNLAARAAWEILQGFSLGSLDSRVQSKDFSLWTES	229	
Db	105	-----GSSG-----VSNTVAAGKQVYNFLELFFDQFPYVNGQ--DFHIAGES	146	

```
QY 230 YGHHYGAFFNHHFYEQNERIANGSVNGVQLNFNSLGI--NGIIDEAIAQPIYPEFA--- 284
D 147 YAGHYIVFASEILSHKDR-----NFNLTSVLINGLTDPLTQNYEPMACGE 195
QY 285 -----VNTYGIKAVNETVNYMKFANOMPGCODLISTCKQNETALADYALCAEATNM 339
D 196 GGEPSVLPSECSAMEDSL-----ERCGLGIESCYDSQ-----SWWSCVPATIIY 239
QY 340 CRDNEVEGYAFAGRGVYDIRHPYDD-----PTPPSYNKFELAKDSVMDAIGVNINYTQS 394
D 240 CNAQLAPYQR-TGRNVYDIRKDCGGNLCYPTLQD-IDDYLNQDYKAEVGAEDVHYES 297
QY 395 -NNDVYAFQOTGDFVWPNFIEDLEEILALPVRVSLIYGDADYICNWFPGQAVSLAANY 453
D 298 CNFIDNRRNLFAGDWMKP-YHTAVTDLLNQDLPILVYAGDKDFICNWLGNKAWTDVLPWK 356
QY 454 QAAQFRS---AGYPLKXNGVEYGETREYGNFSTRVYEAAGHEVYVYQPIASLOLFNRTI 510
D 357 YDEEFASQKVRNWTASITDEVA-GEVKSYPKHTYLRVFNFGHGMVFPDVPENALSNVNEWI 415
QY 511 FG 512
D 416 HG 417

RESULT 2
US-09-901-252-15
; Sequence 15, Application US/09901252
; Patent No. US20020026658A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: Genes Encoding Sinapoylglucose:Malate Sinapoyltransferase and Met
; TITLE OF INVENTION: Use
; FILE REFERENCE: N1422-004
; CURRENT APPLICATION NUMBER: US/09/901,252
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216593
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces
US-09-901-252-15
Query Match 17.4%; Score 517.5; DB 9; Length 421;
Best Local Similarity 29.5%; Pred. No. 6e-40;
Matches 142; Conservative 70; Mismatches 185; Indels 85; Gaps 17;
QY 51 RYKEPGAEGVCETTPGVKSYSGYVDTSP-SHTFFWFFFEARHNPETAPITLWLNKGPGSD 109
D 1 KIKDPKILGI---DPNVTQYTGVLDEDEKHHFFWTFESRNPAPKDPVILWLNKGPGCS 57
QY 110 SLIGLFEELGFCFHVNSFTDDYINPHSNVNSNLFISQPLGVGFVSDDTVGDSINPVTGV 169
D 58 SLTGLFELGFSISGPDLPKPIGNPYSNATVIFLDQPVNVGFYS-----104
QY 170 VENSSFAGVQGRYPTIDATLIDITNLAAEAWEILQGLSLPSLDSRVQSKDFSLWTES 229
D 105 -----GSSG-----VSNTVAAGKDVINELELFFQDFPYVNGKQ--DPIHAGES 146
QY 230 YGHHYGAFFNHHFYEQNERIANGSVNGVQLNFNSLGI--NGIIDEAIAQPIYPEFA--- 284
D 147 YAGHYIVFASEILSHKDR-----NFNLTSVLINGLTDPLTQNYEPMACGE 195
QY 285 -----VNTYGIKAVNETVNYMKFANOMPGCODLISTCKQNETALADYALCAEATNM 339
D 196 GGEPSVLPSECSAMEDSL-----ERCGLGIESCYDSQ-----SWWSCVPATIIY 239
QY 340 CRDNEVEGYAFAGRGVYDIRHPYDD-----PTPPSYNKFELAKDSVMDAIGVNINYTQS 394
D 240 CNAQLAPYQR-TGRNVYDIRKDCGGNLCYPTLQD-IDDYLNQDYKAEVGAEDVHYES 297
QY 395 -NNDVYAFQOTGDFVWPNFIEDLEEILALPVRVSLIYGDADYICNWFPGQAVSLAANY 453
D 298 CNFIDNRRNLFAGDWMKP-YHTAVTDLLNQDLPILVYAGDKDFICNWLGNKAWTDVLPWK 356
QY 454 QAAQFRS---AGYPLKXNGVEYGETREYGNFSTRVYEAAGHEVYVYQPIASLOLFNRTI 510
D 357 YDEEFASQKVRNWTASITDEVA-GEVKSYPKHTYLRVFNFGHGMVFPDVPENALSNVNEWI 415
QY 511 FG 512
D 416 HG 417

RESULT 3
US-10-084-018-3
; Sequence 3, Application US/10084018
; Publication No. US20020160499A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; HAWKINS, PHILLIP R.
; HILLMAN, JENNIFER L.
; LAL, PREETI
; GOLI, SURYA K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,018
; FILING DATE: 25-Feb-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,689A
; FILING DATE: <unknown>
; APPLICATION NUMBER: 08/828,488
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR3DT01
; CLONE: 566993
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-084-018-3
Query Match 12.1%; Score 361; DB 14; Length 476;
Best Local Similarity 26.7%; Pred. No. 4.1e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;
QY 65 PG--VKYSYGY--VDTSPESHHTFFFEARHNPETAPITLWLNKGPGSDSLIGLFEELGP 120
D 67 PGLNKKSYAGPLTYNKTNSNLNFFWFFPAIQPEDAPVILWLNQGGPGSSMXGLFVEHGP 126
```

QY 121 CHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFYSYSDTVGSGINPVTGVWVNSFAGVQG 180
 Db 127 YVWTSNMTLRDRDPFWTTTAXSMLYIDNPNVGTGFSFTDDTHG----- 167
 QY 181 RYPTIDATLDTNLAARAAWEILQGLFSLGSLPSLDSRVQSKDPSLWTSYGGHYGPAPFN 240
 Db 168 -----YAVNEDDVARDLYSALIOFF-----QIFPEYKNDNFYVTGVSAGKYVPAIAH 215
 QY 241 HFYEONERIANGVNGVQLNFNSLGIINGIIDEALQAPYEPFAVNNYTGKAVNETVYN 300
 Db 216 LIHSLN-----PVREKINLNGAIGDGYSDPSIIGGVAEP-----LYQIGLLDEKQK 265
 QY 301 YMKFANQMPNGCODLISTCKQTN---RTALADYALCAEATN---MCRDNVEG---PYAFAG 353
 Db 266 Y--FQKQ-----CHECIEHIRKQNWFEAFIILDKLLDGLTSDPSYFQNVTCGSNYNFI- 318
 QY 354 RGVYDIRHPYDDTPPSYNNKELAKDSVMDAIGVNIYNTQSNNDVY---YAFQQTGDFVWP 411
 Db 319 -----RCTEEDQLYVYKFLSLPEVRQAIHVG-NQTFNDGTIVEKYLREDTVQSVKP 369
 QY 412 NFIEDLEILALPVRVSLIYGDADI-----CNWFGQAVSLAANYSQAAQ-- 457
 Db 370 WLTEIMNN-----YKVLINYGQLDIIVAAALTERSLMGMDWKSQ-----EYKKAQKV 418
 QY 458 ---FRS-----AGYTPLKVNVEYGETREYGNFSTRVYEAGHEVPYYPYQPIASLQLENRTI 510
 Db 419 WKIFKSDSEVAGY-----IRQVGDHFQVIRGGGHILPYDQPLRAFDMINRFI 466
 QY 511 F--GWD 514
 Db 467 YGKGWD 472

RESULT 4

US-09-909-320-164
 ; Sequence 164, Application US/09909320
 ; Patent No. US20020132240A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: ROY, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/909,320
 ; CURRENT FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 60/143,048
 ; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698
 ; PRIOR FILING DATE: 1999-07-26
 ; PRIOR APPLICATION NUMBER: US 60/146,222
 ; PRIOR FILING DATE: 1999-07-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219
 ; PRIOR FILING DATE: 2000-01-05
 ; NUMBER OF SEQ ID NOS: 423
 ; SEQ ID NO 164
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-909-320-164

Query Match

Best Local Similarity 12.1%; Score 359; DB 10; Length 476;
 Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VKSYSGY--VDTSPESTHPTFFFAHNPETAPITLWNGGPGSLSLGLFEELGP 120
 Db 67 PGLNWKSYAGELTVNKNYNSLFFWFFFAQIQEPDAPVVLWLTQGGPGSSMFLGFVEHGP 126
 QY 121 CHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFYSYSDTVGSGINPVTGVWVNSFAGVQG 180
 Db 127 YVWTSNMTLRDRDPFWTTTAXSMLYIDNPNVGTGFSFTDDTHG----- 167
 QY 181 RYPTIDATLDTNLAARAAWEILQGLFSLGSLPSLDSRVQSKDPSLWTSYGGHYGPAPFN 240
 Db 168 -----YAVNEDDVARDLYSALIOFF-----QIFPEYKNDNFYVTGVSAGKYVPAIAH 215
 QY 241 HFYEONERIANGVNGVQLNFNSLGIINGIIDEALQAPYEPFAVNNYTGKAVNETVYN 300
 Db 216 LIHSLN-----PVREKINLNGAIGDGYSDPSIIGGVAEP-----LYQIGLLDEKQK 265
 QY 301 YMKFANQMPNGCODLISTCKQTN---RTALADYALCAEATN---MCRDNVEG---PYAFAG 353
 Db 266 Y--FQKQ-----CHECIEHIRKQNWFEAFIILDKLLDGLTSDPSYFQNVTCGSNYNFI- 318
 QY 354 RGVYDIRHPYDDTPPSYNNKELAKDSVMDAIGVNIYNTQSNNDVY---YAFQQTGDFVWP 411
 Db 319 -----RCTEEDQLYVYKFLSLPEVRQAIHVG-NQTFNDGTIVEKYLREDTVQSVKP 369
 QY 412 NFIEDLEILALPVRVSLIYGDADI-----CNWFGQAVSLAANYSQAAQ-- 457
 Db 370 WLTEIMNN-----YKVLINYGQLDIIVAAALTERSLMGMDWKSQ-----EYKKAQKV 418
 QY 458 ---FRS-----AGYTPLKVNVEYGETREYGNFSTRVYEAGHEVPYYPYQPIASLQLENRTI 510
 Db 419 WKIFKSDSEVAGY-----IRQVGDHFQVIRGGGHILPYDQPLRAFDMINRFI 466

QY	511 F--GWD	514
	:	
Db	467 YGKGWD	472

RESULT 5
US-09-909-088B-164
; Sequence 164, Application US/09909088B
; Patent No. US20020146709A1

```

: Sequence 164, Application US/09309088B
: Patent No. US20020146709A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, A.
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth, J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Macher, Jennie P.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William, I.
:
: TITLE OF INVENTION: Secreted and Transmembrane
:
: TITLE OF INVENTION: Acids Encoding the S
:
: FILE REFERENCE: 10466-14
:
: CURRENT APPLICATION NUMBER: US/09/909,088
: CURRENT FILING DATE: 2001-07-18
:
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
:
: PRIOR APPLICATION NUMBER: US 60/143,048
: PRIOR FILING DATE: 1999-07-07
:
: PRIOR APPLICATION NUMBER: US 60/145,698
: PRIOR FILING DATE: 1999-07-26
:
: PRIOR APPLICATION NUMBER: US 60/146,222
: PRIOR FILING DATE: 1999-07-28
:
: PRIOR APPLICATION NUMBER: PCT/US99/20594
: PRIOR FILING DATE: 1999-09-08
:
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIOR FILING DATE: 1999-09-13
:
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIOR FILING DATE: 1999-09-15
:
: PRIOR APPLICATION NUMBER: PCT/US99/21547
: PRIOR FILING DATE: 1999-09-15
:
: PRIOR APPLICATION NUMBER: PCT/US99/23089
: PRIOR FILING DATE: 1999-10-05
:
: PRIOR APPLICATION NUMBER: PCT/US99/28214
: PRIOR FILING DATE: 1999-11-29
:
: PRIOR APPLICATION NUMBER: PCT/US99/28313
: PRIOR FILING DATE: 1999-11-30
:
: PRIOR APPLICATION NUMBER: PCT/US99/28564
: PRIOR FILING DATE: 1999-12-02
:
: PRIOR APPLICATION NUMBER: PCT/US99/28565
: PRIOR FILING DATE: 1999-12-02
:
: PRIOR APPLICATION NUMBER: PCT/US99/30095
: PRIOR FILING DATE: 1999-12-16
:
: PRIOR APPLICATION NUMBER: PCT/US99/30911
: PRIOR FILING DATE: 1999-12-20
:
: PRIOR APPLICATION NUMBER: PCT/US99/30999
: PRIOR FILING DATE: 1999-12-20
:
: PRIOR APPLICATION NUMBER: PCT/US00/00219

```

RESULT 6
US-09-905-291A-164
; Sequence 164, Application US/09905291A
; Patent No. US2002160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrata, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James

```

; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905/291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR FILING DATE: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR FILING DATE: 1999-07-26
; PRIOR FILING DATE: 1999-07-28
; PRIOR FILING DATE: 1999-07-28
; PRIOR FILING DATE: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR FILING DATE: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR FILING DATE: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR FILING DATE: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR FILING DATE: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR FILING DATE: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR FILING DATE: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR FILING DATE: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR FILING DATE: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-164

Query Match      12.1%; Score 359; DB 10; Length 476;
Best Local Similarity 26.7%; Pred No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VKSYSGY--VDTSPESHTEFFFAHNPETAPITLWNGGPGDLSLGLFEELGP 120
D 67 PGLNMKSYAGFLTKNTYNSLFFWFFPAQIQEPAPVVLWLGPGGSSMFLFVEHGP 126
QY 121 CHVNSTFDYINPHSWNEVSNLLFLSGLVGVGFYSYDVTGDSINPVTGVWVNSSPAGVQG 180
D 127 YVTSNMTLRDRDFPWTTLTSLMLYIDNPGTGFSTDDTHG----- 167
QY 181 RYPTIDATLIDTNLAEEAAWEILQGLSGLPSLDSRVQSKDPSLWTSYGHYGPAPFN 240
D 168 -----YAVNEDEVARDLYSALIGFF-----QIPPEYKNDFFVTGSGYAGKYVPAIAH 215
QY 241 HFYQNERIANGSVNGVQNFNSLGIINGIIDEAQAPYEPFAVNTYGIKAVNETVYN 300
D 216 LIHSIN-----PVREKINLNGAIGDYPESIIIGYAEF-----LYQIGLDEKQKK 265
QY 301 YMKFANQMPNQCQDLISTCKQTN-----RTALADYALCABATN--MCRDNVEG--PYAPAG 353

```

```

Db 266 Y--FOKQ---CHCIEHIRKQNWFEAFIELDKLDGLTSDPSYFQNVTCGSNYNFI-- 318
QY 354 RGVYDIRHPYDDPTPPSYNKFSLAKDSYMDAIGVNIYNTQSNNDVY--YAFQOTGDFWVP 411
D 319 -----RCTEPEDQLYYVAFSLPEVRQAIHVG--NOTFNDGTIVEKYLRDFTVQSVKP 369
QY 412 NFIEDLEILALPVRVSLIYGDADYI-----CNWFGGQAVSLAANYSOAAQ-- 457
D 370 WLTEIMNN-----YKVLIIYNGQLDIIIVAAALTEKSLMCMWKGQSQ-----EYKAEKKV 418
QY 458 ---FRS-----AGYTPKLVNGVEYCEYREYGNFSFTRVREAGHEVYVYQPIASLOLFNRTI 510
D 419 WKIFKSDSEVAGY-----IRQAGDFHQVIRGGGHILPYDQPLRAFDMINRFI 466
QY 511 F--GWD 514
D 467 YGKGWD 472

```

RESULT 7

```

; Sequence 164, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05

```

PRIOR APPLICATION NUMBER: PCT/US99/28214
 PRIOR FILING DATE: 1999-11-29
 PRIOR APPLICATION NUMBER: PCT/US99/28313
 PRIOR FILING DATE: 1999-11-30
 PRIOR APPLICATION NUMBER: PCT/US99/28564
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/28565
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/30095
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: PCT/US99/30911
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US99/30999
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US00/00219
 PRIOR FILING DATE: 2000-01-05
 NUMBER OF SEQ ID NOS: 423
 SEQ ID NO 164
 LENGTH: 476
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-902-853-164

Query Match 12.1%; Score 359; DB 10; Length 476;
 Best Local Similarity 26.7%; Pred. No. 6.3e-25;
 Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

65	PG--VKSYSGY--VDTSPESHTEFFWFHARHNPETAPITLMLNGPGSDSLIGLFEELGP	120
67	PLNWKMSVAGELTVNKNTNSNLFWFPPAQIPEDAPVVLWLGQPGGSSMFLGFVEHGP	126
121	CHVNTEFDYINPHSNWNEVSNLLFLSQLGVGSYSDIVDGSINPVTGVWENSSFAGVQG	180
127	YVVTNNITLRDRDFPWTTLTSLMLYIDNPVGTGFSFTDTHG-----	167
181	RYPTIDATLIDTNNLAEEAAWEILOGLSLPSLDSRVQSKDFSLWTSYGGHYGPAPFN	240
168	-----YAVNEDVDARDLYSALIQFF-----QIFPEYKNNDFYVTGESYAGKYVPATAH	215
241	HFYQNERIANGSVQVNFNSLGIINGIIDEALQAPYPPFAVNNYGIKAVNETVYN	300
216	LTHSLN-----PVREKINLNGAIGDYSDPESIIGGYAEF-----LYOIGLLDEKQKK	265
301	YNKFANQMPGCGDILSTCKQTN--RTALADYALCAEATN--MCRDNVEG--PYAFAG	353
266	Y--FOKQ-----CHECHIRKQNWFAFELDKLDDLTSPSYFQNTVGCNSYNYFL--	318
354	RGVYDIRHPYDDTPPSYNNKFLAKDSVMDAIGVINITYTQSNNDVY--YAFQOTGDFWFP	411
319	-----RCTEPEDQLYVVKFLSLPEVRQAIHVG--NQTENDGTIVEKYLREDTVQSVKP	369
412	NFTEDLEELALPVRSLYINGDADYI-----CNWFGQVAVLAANYSOAAQ--	457
370	WLTEIMNN-----YKVLIVNGQLDIIVAALTERSLMGMDWKGSO-----EYKKAQKV	418
458	---FRS-----AGYTPLKXVNGVEYGETREYGNFSFTRVYEGAEHVPYQPIASQLFNRTI	510
419	WKLFKSDSEVAGY-----IRQAGDFEQVIIRGGGHILPYDQPLRAFDMINRFI	466
511	F--GWD 514	
467	YGRGWD 472	

RESULT 8
 US-09-907-824-164
 Sequence 164, Application US/09907824

APPLICANT: Baton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, A.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth, J.
 APPLICANT: KJavin, Ivar J.
 APPLICANT: Mather, Jennie P.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William, I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same
 FILE REFERENCE: 10466-14
 CURRENT APPLICATION NUMBER: US/09/907,824
 CURRENT FILING DATE: 2001-07-17
 PRIOR APPLICATION NUMBER: 09/665,350
 PRIOR FILING DATE: 2000-09-18
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: US 60/143,048
 PRIOR FILING DATE: 1999-07-07
 PRIOR APPLICATION NUMBER: US 60/145,698
 PRIOR FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: US 60/146,222
 PRIOR FILING DATE: 1999-07-28
 PRIOR APPLICATION NUMBER: PCT/US99/20594
 PRIOR FILING DATE: 1999-09-08
 PRIOR APPLICATION NUMBER: PCT/US99/20944
 PRIOR FILING DATE: 1999-09-13
 PRIOR APPLICATION NUMBER: PCT/US99/21090
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/21547
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/23089
 PRIOR FILING DATE: 1999-10-05
 PRIOR APPLICATION NUMBER: PCT/US99/28214
 PRIOR FILING DATE: 1999-11-29
 PRIOR APPLICATION NUMBER: PCT/US99/28313
 PRIOR FILING DATE: 1999-11-30
 PRIOR APPLICATION NUMBER: PCT/US99/28564
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/28565
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/30095
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: PCT/US99/30911
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US99/30999
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US00/00219
 PRIOR FILING DATE: 2000-01-05
 NUMBER OF SEQ ID NOS: 423
 SEQ ID NO 164
 LENGTH: 476
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-907-824-164

```

RESULT 8
US-09-907-824-164
: Sequence 164, Application US/09907824
: Publication NO. US20020197671A1
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Botstein, David
: APPLICANT: Desnovers, Luc
: LENGTH: 476
: TYPE: PRT
: ORGANISM: Homo Sapien
US-09-907-824-164
Query Match 12.1%; Score 359; DB 10; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

```


Query Match 12.1%; Score 359; DB 11; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VXSYSY--VDTSPESHTFFFEARHNPETAPITLWLNKGPSDSLGLFELGP 120
DB 67 PGLNMSYAGFLVNTKYNLFFWPPFAIQIPEDAPVLWLQGGPGSSMFLGVEHGP 126
QY 121 CHVNSTFDDYINPHSNVSNLLFLSQPLGVGYSYSDTVDGSINFTGVVSNSSFAGVQG 180
DB 127 YVVTNNMTLRDRDPPTTILSMLYIDNPVGTFSTDDTHG----- 167
QY 181 RYPTIDATLDTNLAEEAAWEILQGLSLPSLDSRVQSKDFSLWTSYGHYGPFFN 240
DB 168 -----YVNEDDVARDLYSALIQFF-----QIFPEYKKNDFVVTGESYAGKVPAIAH 215
QY 241 HFVEQNERIANGSVNGVQLNFSNGIINGIIDEAIAQIAPYPEFVAVNNTYGIKAVNETVYN 300
DB 216 LIHSLN-----PVREVKINLNGIAGDYSDPESIIGGYAEF-----LYQIGLLDEKQK 265
QY 301 YMKFANQMPNGCQDLISTCKQTN---RTALADYALCAEATN--MCRDNVEG--PYAFAG 353
DB 266 Y--FQKQ---CHECIEHIRKQNWFEAFELDKLDDGLTSDPSYFQVNTGCSNYNPL- 318
QY 354 RGVDYIRHPYDDPTPPSYNKFSLAKDSYMDAIGVNNYNTQSNNDVY--YAFQQTGDFVMP 411
DB 319 -----RTEPEDQLYVVKFSLPEVRQAIHVG--NOTFNDGTIVEKYLREDTVQSVKP 369
QY 412 NFIEDLEILALPVRSYLIYGDADYI-----CNWFGGQAVSLAANYSOAQ-- 457
DB 370 WITEIMNN-----YKVLINGQLDIIIVAAALTERSLMGMDWKGSQ-----EYKAEKKV 418
QY 458 ---FRS-----AGYTPLKNGVEYGETREYGNFSTRVYEAGHEVPYPIASLQLFNRTI 510
DB 419 WKIFKSDSEVAGY-----IROAGDFHVIIRGGHILEFDQPLAFDMINRPI 466
QY 511 P--GWD 514
DB 467 YGKGWD 472

RESULT 11
US-09-906-742-164
; Sequence 164, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-164

```
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,742
CURRENT FILING DATE: 2001-07-16
PRIOR FILING DATE: 2000-09-18
PRIOR FILING DATE: 2000-09-18
PRIOR FILING DATE: 2000-02-22
PRIOR FILING DATE: 1999-07-07
PRIOR FILING DATE: 1999-07-26
PRIOR FILING DATE: 1999-07-28
PRIOR FILING DATE: 1999-07-28
PRIOR FILING DATE: 1999-09-08
PRIOR FILING DATE: 1999-09-13
PRIOR FILING DATE: 1999-09-15
PRIOR FILING DATE: 1999-11-29
PRIOR FILING DATE: 1999-11-30
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-16
PRIOR FILING DATE: 1999-12-20
PRIOR FILING DATE: 1999-12-20
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 164
LENGTH: 476
TYPE: PRT
ORGANISM: Homo Sapien
US-09-906-742-164

Query Match
Best Local Similarity 12.1%; Score 359; DB 11; Length 476;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

65 PG--VKSYSGY--VDTSPESHFFWFFFEARHNPETAPITLWNGFGSDSLGLFEEELGP 120
67 PGLANKSYAGELTVNKTYSNLFWFFFAIQIPEDAPVVLWLGQFGGSSMGLFVEHGP 126
121 CHVNSTEDYINPHSWNEVSNLFLSQPLGVGFSYSDTVGSGINPVTGVVENSFAGVQG 180
127 YVTSNMTLRDPPFTWTTLSMLYDNPVGTGFSFTDTHG-----167
181 RYPTIDAFTIDTTLAAAEWELLGSLGSLPSLDSRVQSKDFSLWTSYGHYGPAPFN 240
168 -----YAVNEDDVARDLYSALQFF-----QIPPEYKNDFFVTGSGYAGKYVPAIAH 215
241 HFYQNEIRIANGSVNGVQNFNSLGIINGIIDEAIAQAYYPEFAVNNYTGKAVNETVYN 300
216 LIHSIN-----PVREVKNLNGIAIGDYSDPESIIIGYAEF-----IYQIGLDEKQKK 265
301 YMKFANQPNQCODLSTCKQTN---RTALADYALCAEATN---MCRDNVEG---PYAPAG 353
266 Y--FOKQ---CHECIEHQRQNWFPEAILEDKLLDGLTSDPSYFQNVTCSSNYNPL- 318
354 RGVYDIRHPYDPTPPSYNKFELAKDSYMDAIGNYNINVTQSNNDV---YAFQQTGDFWVP 411
319 -----RCTEPEDQLYVFKLSLPEVRQAIHVG-NQTFNDGTIVKYLREDTVQSVKP 369

FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,742
CURRENT FILING DATE: 2001-07-16
PRIOR FILING DATE: 2000-09-18
PRIOR FILING DATE: 2000-09-18
PRIOR FILING DATE: 2000-02-22
PRIOR FILING DATE: 1999-07-07
PRIOR FILING DATE: 1999-07-26
PRIOR FILING DATE: 1999-07-28
PRIOR FILING DATE: 1999-07-28
PRIOR FILING DATE: 1999-09-08
PRIOR FILING DATE: 1999-09-13
PRIOR FILING DATE: 1999-09-15
PRIOR FILING DATE: 1999-11-29
PRIOR FILING DATE: 1999-11-30
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-02
PRIOR FILING DATE: 1999-12-16
PRIOR FILING DATE: 1999-12-20
PRIOR FILING DATE: 1999-12-20
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 164
LENGTH: 476
TYPE: PRT
ORGANISM: Homo Sapien
US-09-906-742-164

US-09-906-838-164
Sequence 164, Application US/09906838
Publication No. US20030027143A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,838
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30

RESULT 12
US-09-906-838-164
Sequence 164, Application US/09906838
Publication No. US20030027143A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,838
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-906-838-164

Query Match      12.1%; Score 359; DB 11; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

Qy 65 PG--VKSYGY--VDTSPESHTEFFFEARHNPTAPITLWNGGPGSDSLIGLFEELGP 120
Db 67 PGLNMKSYAGFLTVNKTNSNLFWFFFPAQIQEPADAPVVLWLOGGPGSSMFGFLVBEHGP 126
Qy 121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGVWENSSPAGVQG 180
Db 127 YVVTSMTLRDRDPFWITTLNMLYIDNPVGTGFSFTDTHG----- 167
Qy 181 RYPTIDATLIDTNLAEEAAWEILQGLSLPSLDSRVQSKDFSLWTSYGGHYGPAPFN 240
Db 168 -----YAVNEDDVARDLYSALIOFF-----QIFPEYKKNDFYVTGSGYACKYVPAIAH 215
Qy 241 HFEQNERIANGSVNGVQNFNSLGIINGIDEAQPYPPFVANNYGIKAVNETVYN 300
Db 216 LIHSLN-----PVREVKINLNGIAGDGYSDPSIIGYAEF-----LYQIGLLDEKQKK 265
Qy 301 YMKPANQMPNGQDILSTCKQTN---RTALADYALCAEATN---MCRDNRVGE--PYAFAG 353
Db 266 Y--FOKQ---CHECHIEIRKQNWFEPAFEILDKLDGDLTSPSFQNTGSGNYNLF- 318
Qy 354 RGWYDIRHPDPTPSYKFKIADKSDVMDAIGVNNYITQSNNDVY--YAFQGTGFVWP 411
Db 319 -----RCTSPEDQLYYKFLSLPEVRQAIVHG-NQTFNDGTIVEKYLREDTVQSVKP 369
Qy 412 NFIEDLEEILALPVRVSLYGDADYI-----CNWFGQAVSLAANYSQAAQ-- 457
Db 370 WLTEIMNN-----YKVLIVNGQLDIIVAAALTERSLMGMDWKGSQ-----EYKKAERKV 418
Qy 458 ---FRS---AGVTPKLVNGVEYGETREYGNFSTFVZBAGHEVPYVYQPIASQLPFRIT 510
Db 419 WKIFKSDSEVAGY-----IRQAGDFHQVIRGGHILPYDQPLRAEDMINRFI 466
Qy 511 F--GWD 514
Db 467 YGKGWD 472

RESULT 13
US-09-907-613-164
; Sequence 164, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,613
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-613-164

Query Match      12.1%; Score 359; DB 11; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

Qy 65 PG--VKSYGY--VDTSPESHTEFFFEARHNPTAPITLWNGGPGSDSLIGLFEELGP 120
Db 67 PGLNMKSYAGFLTVNKTNSNLFWFFFPAQIQEPADAPVVLWLOGGPGSSMFGFLVBEHGP 126
Qy 121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGVWENSSPAGVQG 180
Db 127 YVVTSMTLRDRDPFWITTLNMLYIDNPVGTGFSFTDTHG----- 167
Qy 181 RYPTIDATLIDTNLAEEAAWEILQGLSLPSLDSRVQSKDFSLWTSYGGHYGPAPFN 240
Db 168 -----YAVNEDDVARDLYSALIOFF-----QIFPEYKKNDFYVTGSGYACKYVPAIAH 215
Qy 241 HFEQNERIANGSVNGVQNFNSLGIINGIDEAQPYPPFVANNYGIKAVNETVYN 300
Db 216 LIHSLN-----PVREVKINLNGIAGDGYSDPSIIGYAEF-----LYQIGLLDEKQKK 265
Qy 301 YMKPANQMPNGQDILSTCKQTN---RTALADYALCAEATN---MCRDNRVGE--PYAFAG 353
Db 266 Y--FOKQ---CHECHIEIRKQNWFEPAFEILDKLDGDLTSPSFQNTGSGNYNLF- 318
Qy 354 RGWYDIRHPDPTPSYKFKIADKSDVMDAIGVNNYITQSNNDVY--YAFQGTGFVWP 411
Db 319 -----RCTSPEDQLYYKFLSLPEVRQAIVHG-NQTFNDGTIVEKYLREDTVQSVKP 369
Qy 412 NFIEDLEEILALPVRVSLYGDADYI-----CNWFGQAVSLAANYSQAAQ-- 457
Db 370 WLTEIMNN-----YKVLIVNGQLDIIVAAALTERSLMGMDWKGSQ-----EYKKAERKV 418
Qy 458 ---FRS---AGVTPKLVNGVEYGETREYGNFSTFVZBAGHEVPYVYQPIASQLPFRIT 510
Db 419 WKIFKSDSEVAGY-----IRQAGDFHQVIRGGHILPYDQPLRAEDMINRFI 466
Qy 511 F--GWD 514
Db 467 YGKGWD 472
```

```

Db      127 YVVTSNMTLRDRFPWTTTSLMYLDNPVGTGFSFTDDTHG----- 167
Qy      181 RYPTIDATLIDTNNLAEEAAWEILQGLFSLDSRVQSKDFSLWTESYGHYGPAFEN 240
Db      168 -----YAVNEDDVDARDLYSALIOFF-----QIPPEYKNNDFYVTGESYAGKYVPAINH 215
Qy      241 HFYEQNERIANGSVNGVLNFNSLGIIINGIIDEAQAPYPPEFAVNNTVGIAKNETVYN 300
Db      216 LIHSLN-----PVREVKINLNGIAIGDYSPEFIIGGYAEF---LYQIGLLDEKOKK 265
Qy      301 YMKFANQMPPENGQQDLISTCKOTN---RVALADYALCAEATN--MCRDNVEG--PYVAFAG 353
Db      266 Y--PQKQ----CHECIEHRQNWFEPFEILDKLGDGLTSDPSYFQNVGTGCSYNYFL- 318
Qy      354 RGVDYDIRHPDYDOPTPPSYNNKFLAKDSVMYDAIGNVINYTQSNDVY--YAFQOTGDFVWP 411
Db      319 -----RCTEPEDQLYYVWFSLUPEVRQAHHVG-NQTFNDDGTIVEXYLREDTVQSVKP 369
Qy      412 NFIEDLEILALPVRVSLIYGDADI-----CNWFGGAVSLSAANYSAQAQ--- 457
Db      370 WLTEIMNN-----YKVLIYNGQLDIIVAAATERSLMGMWDKWSQ-----EYKKAERKV 418
Qy      458 ---PRS-----AGYTPLKVNGVEYGETREYGNFSTRVYEAGHEVPYQPIASLOLFNRTI 510
Db      419 WKIFKSDSEVAGY-----IRGAGDFHQVIIRGGGHILPYDOPDLRAFAMINFRI 466
Qy      511 F--GWD 514
Db      467 YKGKWD 472

```

```

RESULT 14
US-09-907-942-164
// Sequence 164, Application US/0907942
// Publication No. US20030027146A1
// GENERAL INFORMATION:
// APPLICANT: Genentech, Inc
// APPLICANT: Ashkenazi, Avi
// APPLICANT: Botstein, David
// APPLICANT: Desnoyers, Luc
// APPLICANT: Eaton, Dan L.
// APPLICANT: Ferrara, Napoleone
// APPLICANT: Filvaroff, Ellen
// APPLICANT: Fong, Sherman
// APPLICANT: Gao, Wei-Qiang
// APPLICANT: Gerber, Hanspeter
// APPLICANT: Gerritsen, Mary E.
// APPLICANT: Goddard, A.
// APPLICANT: Godowski, Paul J.
// APPLICANT: Grimaldi, Christopher J.
// APPLICANT: Gurney, Austin L.
// APPLICANT: Hillan, Kenneth, J.
// APPLICANT: Kljavin, Ivar J.
// APPLICANT: Mather, Jennie P.
// APPLICANT: Pan, James
// APPLICANT: Paoni, Nicholas F.
// APPLICANT: Roy, Margaret Ann
// APPLICANT: Stewart, Timothy A.
// APPLICANT: Tumas, Daniel
// APPLICANT: Williams, P. Mickey
// APPLICANT: Wood, William, I.
// TITLE OF INVENTION: Secreted and Transmitted
// TITLE OF INVENTION: Acids Encoding tRNA
// FILE REFERENCE: 10466-14
// CURRENT APPLICATION NUMBER: US/09/907942
// CURRENT FILING DATE: 2002-01-22
// PRIOR APPLICATION NUMBER: PCT/US00/0422
// PRIOR FILING DATE: 2000-02-22
// PRIOR APPLICATION NUMBER: US 60/143,000
// PRIOR FILING DATE: 1999-07-07
// PRIOR APPLICATION NUMBER: US 60/145,616
// PRIOR FILING DATE: 1999-07-26
// PRIOR APPLICATION NUMBER: US 60/146,216

```

; PRIOR FILING DATE: 1999-07-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219
 ; PRIOR FILING DATE: 2000-01-05
 ; NUMBER OF SEQ ID NOS: 423
 ; SEQ ID NO 164
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-907-942-164

Query Match	12.1%	Score 359;	DB 11;	Length 476;
Best Local Similarity	26.7%;	Pred No. 6.3e-25;		
Matches 130;	Conservative 62;	Mismatches 178;	Indels 116;	Gaps 21;
Qy 65	PG--VKSYSGY--VDTSPESHFTFWFPEARHNPTAPITLWLNGPGSDSLIGLFEELGP	120		
Db 67	PLGNKYSAGELTVKNTYNSLNFEEFPAQIQPEDAPVLWLQGGPGGSMFGLFVEHGP	126		
Qy 121	CHVNSTFDDYINPHSMNEVSNLFLSLOPLGVGSYSDTVDGSINPTVGWVENSFAGVQG	180		
Db 127	VYVTSNLTDRDRDPWTTTLLSMLYIDNPVGTSFSTDDTHG	167		
Qy 181	RYPTIDATLIDTNLAAEAAWEILQGLSLGPSIDSRVQKDFSLWTESYGGHYGPAFFN	240		
Db 168	YAVNEDDVARLDYSA LIQFF-----QIFPEYKANDFVVTGESVAGKVVPAAIAH	215		
Qy 241	HFEQNEPRIANGSVNGVQLMNSLGIINGIIDIAPAIQAPYYPEFVANNYTYGIKAVNETVYN	300		
Db 216	LIHSLN-----PVREVKINLNGIAGDGYSDPESIIGVAAEF-----LYQIGLLDEKQK	265		
Qy 301	YMKFANQMPNGCODLSTCKQTN--RTALADVALCAEATN--MCRDNVEG--PYAFAG	353		
Db 266	Y--FQKQ-----CHECTEHIRKQWFEAFELDKLLDGDLTSDSFYQNVTCGSNYVNF--	318		
Qy 354	RGVYDIRHPYDDTPPESYNNKFLAKDSVMDAIGWLNINYTOSNNDVY--YAFQQTGDFVWP	411		
Db 319	-----RCTEPEDQLYYVKFLSLPEVRQAIIHVG--NOTFNDGTIVEKYVUREDTVQSVKP	369		
Qy 412	NFTIEDLEETIALPVRVSLYIGDADYI-----CNWFGQAVSLAANYSQAQ--	457		
Db 370	WLTEIMNN-----YKVLIIYNGQLDITVAALATERSLWGMGMDKGSQ-----EYKKAQKV	418		
Qy 458	---FRS-----AGYTPLKVGNGVEGETREYGNFSFTFVYEAGHEVPYQPIASLQLPNRTI	510		
Db 419	WKIFKSDSEVAGY-----IROAGDFHQVIIRGGGHILPYDQPLRAFDMINRTI	466		
Qy 511	F--GND 514			
Db 467	YKGGWD 472			

Mon Nov 24 13:41:12 2003

Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;	
QY	65 PG--VKSYSY--VDTSPESHTRFWRPPEARHNPETAPITLWNGPGQSDSLIGLFEELGP 120
Db	67 PGLNKSAYAGFLTVNKTYSNLFVFPFPAQLQPEDAEVVLWLOGPGGSSMFLFVEHGP 126
QY	121 CHVNSTEDDYINPHSMNEVSNLLFLSQPLGVGSYSYDTPVGSINPVTGVVENSFAVQOG 180
Db	127 YVVTSNMTRDRDPFWTTLSMLYIDNPVGTGFSFTDDTHG----- 167
QY	181 RYPTIDATLIDTNNLAABAAWEILOGLSLPSLDSRVQSKDFSLWTESGHHYGPAPFN 240
Db	168 -----YAVNEDDVARLYSALIQFF-----QIFPEYKNNDFVVTGESYAGKVPAIAH 215
QY	241 HFYEONERLANGSVNGVQLNENSLGIINGIIDEALQAPYYPPEFAVNNTYGKAVNETVYN 300
Db	216 LIHSLN-----PVREVKINLNGIAIGDGYSDPESITGGYAEF-----LYQIGLLDERKQK 265
QY	301 YMKFANQMPNGCQDLISTCKQTN---RTALADYALCAEATN--MCRDNVEG--PYIYAFAG 353
Db	266 Y--FQKQ---CHECIEHIRKQNWFEAFELDKLDDGLTSDPSYFQNVTCNSYNYFL- 318
QY	354 RGVYDIRHPYDDPTPPSYNNKFLAKDSVMDAIGNINVTQSNNDY--YAFQOTGDFVWP 411
Db	319 -----RCTEPEDQLYVVRFLSLPEVRQAIHVG-NQTFNDGTIVEKYLREDITVQSVKP 369
QY	412 NFIEDLEILALPVRVSLIYGDADVI-----CNWFGGOAVSLAANYSOAAQ-- 457
Db	370 WLTEIMNN-----YKVLINYGQLDITIVAAALTERSLMGMDWKGSO-----EYKKAEEKV 418
QY	458 ---FRS---AGYTPLKVGVEYGETREYGNFSTFRVYEAGHEVPYQDIASLQLENRTI 510
Db	419 WKIFKSDSEVAGY-----IRQAGDPHQVIRGGGHILPYDQPLRAFDMINRTI 466
QY	511 F--GWD 514
Db	467 YGKGWD 472

Search completed: November 21, 2003, 18:02:55
Job time : 37 secs

US-09-796-753-40
; Sequence 40, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 40
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-40

Query Match 12.1%; Score 359; DB 11; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:58:46 ; Search time 21 Seconds
(without alignments)
1118.216 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 2979
Sequence: 1 MRGYEFUSVLPVLAASWALP.....HTQSSVPLPTATSMSSVGMA 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCUTS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2975	99.9	554	3	US-08-943-714-2
2	894	30.0	423	3	US-08-943-714-9
3	630	21.1	557	1	US-08-309-341-2
4	630	21.1	557	1	US-08-608-267-2
5	630	21.1	557	1	US-08-608-452-2
6	630	21.1	557	1	US-08-608-224-2
7	630	21.1	557	2	US-08-967-149-2
8	623	20.9	557	1	US-08-309-341-4
9	623	20.9	557	1	US-08-608-267-4
10	623	20.9	557	1	US-08-608-452-4
11	623	20.9	557	1	US-08-967-149-4
12	623	20.9	557	2	US-08-608-224-4
13	568.5	19.1	491	1	US-09-640-305-4
14	568.5	19.1	491	1	US-08-360-673-4
15	523.5	17.6	532	2	US-08-893-324-33
16	523.5	17.6	532	3	US-08-329-892B-33
17	513.5	17.2	421	2	US-08-807-263-4
18	361	12.1	476	2	US-08-828-488-3
19	361	12.1	476	4	US-09-299-689A-3
20	348.5	11.7	477	2	US-08-828-488-1
21	348.5	11.7	477	4	US-09-299-689A-1
22	336	11.3	480	2	US-08-828-488-8
23	336	11.3	480	4	US-09-299-689A-8
24	336	11.3	480	4	US-09-702-705-336
25	336	11.3	480	4	US-09-736-457-336
26	312.5	10.5	471	2	US-08-828-488-7
27	312.5	10.5	471	4	US-09-299-689A-7

28	267	9.0	523	3	US-08-943-714-11	Sequence 11, Appl
29	246.5	8.3	446	1	US-08-665-966-10	Sequence 10, Appl
30	246.5	8.3	446	3	US-09-041-780-10	Sequence 10, Appl
31	208	7.0	530	3	US-08-943-714-12	Sequence 12, Appl
32	204.5	6.9	481	3	US-08-943-714-10	Sequence 10, Appl
33	193	6.5	35	3	US-08-943-714-5	Sequence 5, Appl
34	172	5.8	179	1	US-08-665-966-8	Sequence 8, Appl
35	172	5.8	179	3	US-09-041-780-8	Sequence 8, Appl
36	169.5	5.7	351	2	US-08-828-488-5	Sequence 5, Appl
37	169.5	5.7	351	4	US-09-299-689A-5	Sequence 5, Appl
38	142	4.8	29	3	US-08-943-714-6	Sequence 6, Appl
39	131	4.4	618	3	US-09-199-290-34	Sequence 34, Appl
40	127.5	4.3	771	4	US-09-634-238-229	Sequence 229, App
41	127.5	4.3	771	4	US-09-634-238-262	Sequence 262, App
42	126	4.2	385	2	US-08-387-942C-24	Sequence 24, Appl
43	126	4.2	1403	2	US-08-387-942C-3	Sequence 3, Appl
44	120	4.0	385	2	US-08-387-942C-27	Sequence 27, Appl
45	120	4.0	872	2	US-08-387-942C-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dambmann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-714-2

Query Match

99.9% Score 2975; DB 3; Length 554;

```
Best Local Similarity 100.0%; Pred. No. 4.4e-263; Indels 0; Gaps 0;
Matches 554; Conservative 0; Mismatches 0;

QY 1 MRGYEFLSVPLVAASWALPGSTPASVGRQLPKNPTGVKLTITANNVTIRYKEPGAEGV 60
Db 1 MRGYEFLSVPLVAASWALPGSTPASVGRQLPKNPTGVKLTITANNVTIRYKEPGAEGV 60
QY 61 CETTPGVKSYGYVDTSPESHTEFFWFPEARHNPETAPITLWNGPGSDSLIGLFEELGP 120
Db 61 CETTPGVKSYGYVDTSPESHTEFFWFPEARHNPETAPITLWNGPGSDSLIGLFEELGP 120
QY 121 CHVNSTEDDYINPHSMNEVSNLLFLSQPLGVGFYSYDVTGDSINPVTGVWVNSSFAGVQG 180
Db 121 CHVNSTEDDYINPHSMNEVSNLLFLSQPLGVGFYSYDVTGDSINPVTGVWVNSSFAGVQG 180
QY 181 RYPTIDATLIDTTNLAAEAAWEILQGLSLPSLDSRVQSKDFSLWTESYGGHYGPAPFN 240
Db 181 RYPTIDATLIDTTNLAAEAAWEILQGLSLPSLDSRVQSKDFSLWTESYGGHYGPAPFN 240
QY 241 HFYQNERIANGSVNGVQNFNSLGIINGIIDEAIOAPYPEFAVNNYTGKAVNETVYN 300
Db 241 HFYQNERIANGSVNGVQNFNSLGIINGIIDEAIOAPYPEFAVNNYTGKAVNETVYN 300
QY 301 YMKFANQMPNGQDILSTCKQINRTALADYALCAEATNMCRDNVEGPIYAFAGRGVYDIR 360
Db 301 YMKFANQMPNGQDILSTCKQINRTALADYALCAEATNMCRDNVEGPIYAFAGRGVYDIR 360
QY 361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNIYNTQSNNDVYVAFQOTGDFWPNFIEDLEI 420
Db 361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNIYNTQSNNDVYVAFQOTGDFWPNFIEDLEI 420
QY 421 LALPVRVSLIYGADYICNWFQGVQVSLAANYQAAQFSAQYTPKLVNGVYGETREYG 480
Db 421 LALPVRVSLIYGADYICNWFQGVQVSLAANYQAAQFSAQYTPKLVNGVYGETREYG 480
QY 481 NFSFTRVYBAGHEVPYQPIASLQLENRTIFGWDIAEGOKKIWPSYKTINGTATATHTQSS 540
Db 481 NFSFTRVYBAGHEVPYQPIASLQLENRTIFGWDIAEGOKKIWPSYKTINGTATATHTQSS 540
QY 541 VPLPTATSMSSVGM 554
Db 541 VPLPTATSMSSVGM 554

RESULT 2
US-08-943-714-9
; Sequence 9, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berka, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dambmann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASCSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
```

```
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-943-714-9

Query Match 30.0%; Score 894; DB 3; Length 423;
Best Local Similarity 40.1%; Pred. No. 3.5e-73;
Matches 186; Conservative 66; Mismatches 160; Indels 52; Gaps 8;

QY 59 GVCETTPGVKSYGYVDTSPESHTEFFWFPEARHNPETAPITLWNGPGSDSLIGLFEEL 118
Db 6 GICETTPGVNYSYGYLSVGSNNMWFPEARNPQQAPLAWFNGPGSCSMIGLFOFN 65
QY 119 GPCH-VNSTEDDYINPHSMNEVSNLLFLSQPLGVGFYSYDVTGDSINPVTGVWVNSSFAG 177
Db 66 GPCHFWNGDSTPSLNSNNYNNMIYIDQPLGVGFYSY- 104
QY 178 VQGRYPTIDATLIDTTNLAAEAAWEILQGLSLPSLDSRVQSKDFSLWTESYGGHYGPA 237
Db 105 -----TDDVT---STVTAAPVVMNLQAFYAQRPEYESR---DFAIFTESYGGHYGPE 151
QY 238 PFNHFYQNERIANGSVNGVQNFNSLGIINGIIDEAIOAPYPEFAVNNYTGKAVNET 297
Db 152 FASYIEQQMAAIKAGSVTGQNVNIVALGVNNGWIDSTIQEKAYIDFSYNNYSQQIIDST 211
QY 298 VTNMKFANQMPNGQDILSTCKQINRTALADYALCAEATNMCRDNVEGPIYAFAGRGVY 357
Db 212 RDSLLDAYN---NOCLPALQOCSQSGST-----SDCTNADSVCIQNIIEGPISSSGDFDVY 263
QY 358 DIRHPYDDPTPPSYNKKFLAKDSVMDAIGVNIYNTQSNNDVYVAFQOTGDFWPNFIEDL 417
Db 264 DIREPSNDPYPKTYSTYLSLDSFTVVKAIARTNYQECNPGYKFKASTGDNP-RSFLSTL 322
QY 418 EEIALPVRVSLIYGADYICNWFQGVQVSLAANYQAAQFSAQYTPKLVNGVYGETR 477
Db 323 SSVVQSGINVLVWAGDADWICNWLGNIEVANAVDFPNAQFSAALDAPYTVNGVEKGQFK 382
QY 478 EYGNFSFTRVYBAGHEVPYQPIASLQLENRTIFGWDIAEGOKK 521
Db 383 TVDNFSFLKYVGAGHEVPYQPIASLQLENRTIFGWDIAEGOKK 418

RESULT 3
US-08-309-341-2
; Sequence 2, Application US/08309341
; Patent No. 5594119
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5594119o No. 5594119disk of No. 5594119th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/309,341
; FILING DATE: 16-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-309-341-2

Query Match 21.1%; Score 630; DB 1; Length 557;
Best Local Similarity 31.8%; Pred. No. 7.2e-49;
Matches 154; Conservative 81; Mismatches 174; Indels 76; Gaps 17;

QY 45 ANNVTIRYKEPGAGVCETTPGVKSYGYVDTSP-ESHPTFPWFPEARHNPTAPITLWLN 103
DB 132 AYDRVKRTDPSGLI---DFGQKYTYGLDDNENDKHLFWFFESRNDPENDPVLWLN 188
QY 104 GPGSDSLIGLFEELGPGCHVNSTPDDYINPHSMNEVSNLLFSLQPLGVGFSYSTVDGSI 163
DB 189 GPGCSSLTGLFMELGPSINKKIQPVNDYAMNSASVIFLDQPVNVGYSNSA--- 244
QY 164 NPVTGVVENSFAGVGQRYPTIDATLIDTTLNLAEEAAWEILQGLSLGSLDSRVQSKDF 223
DB 245 -----VSDTVAAQKQVYALLTLFFKQFP-----EYAKQDF 274
QY 224 SLWTESGGHYGPAFFNFHYEQNERIANGSVNGVQLNFSLGHINGIIDEAIQAPYVPF 283
DB 275 HIAGESYAGHYIPVFASEILLSKKR-----NINLQSVLIGNGLTDCYTYQYVYRPM 325
QY 284 AVNNTYGIKAVNEPVYNYMKFANCMENGCODLSTCKQTNRTLADYALCAEATNMCRDN 343
DB 326 ACGDGGYPAVLDRSSCCSMD--NALPR-COSMIETESYSSESA-----WVCVPASIIYCNA 377
QY 344 VEGPYAFAGRGYVDIRHPYDDP-----TPPSYNNKFLAKDSYMDAIGVNNIN-YTOSNND 397
DB 378 LLAPYQR-TQGNVYDVRGKCEDSSNLCSYSAWGVSVDYLNKPFVIEAVGAEVNGYDSCNFD 436
QY 398 VYVAFQQTGDFWVP--NFTLEBIILALPVRSVLIY-GDADYICNFWGQAVSLAANYSQ 454
DB 437 INRNFLFHGDMMKPYHRLVFGLLF--QIPV--LIIVAGDAFICNLWGLNKAWEALEWFG 491
QY 455 AAQFRSAGYVTPLKV-----NGVEYGETREYCNFSFTRVYEAGHEVYYPYQIASLQFNRT 509
DB 492 QREYASAELEDLVINEHTGKIGQVKGHGNFTFMRLYGGGHMVMPMDQPESSELEFFNW 551
QY 510 IFG-W 513
DB 552 LGGEW 556

RESULT 4
US-08-608-267-2
; Sequence 2, Application US/08608267
; Patent No. 5688663
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann

```

TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5688663o No. 5688663disk of No. 5688663th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6400
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/608/267
 FILING DATE: 28-FEB-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/309,341
 FILING DATE: 20-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Lowney, Karen A.
 REGISTRATION NUMBER: 31,274
 REFERENCE/DOCKET NUMBER: 4247.000-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 557 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Aspergillus Nigier
 US-08-608-267-2

Query Match	21.1%	Score 630;	DB 1;	Length 557;	
Best Local Similarity	31.8%;	Pred. No. 7.2e-49;			
Matches	154;	Conservative	81;	Mismatches	174;
				Indels	76;
				Gaps	17;
QY	45	ANNVTIRYKPGAGVCETPGUKSYGYVDTSP-ESH	TFWFFEARHNPETAPITLWLN	103	
Db	132	AYDLRVKTTDFGSLGI---DPGKQYTG	LYLDDNDKELFYWFFESNDPENDPWLWLN	188	
QY	104	GGPGSDSLGLFBEELGCHVNSTFDDYINPHS	MNEVSNLLPLSPLGCGSYGSYSTDVDSI	163	
Db	189	GGPGCSLTLGLFMELGPSSINKKIQPVYND	YANNASNAVIFLOQPVNVGYSYSNSA----	244	
QY	164	NPVTGVVENSFPAGVQGRYPTIDATLIDT	NLAAEAAWEILQGLSLGSLDSRVQSKDF	223	
Db	245	-----VSDTVAAGKDVIALLTLTFKQFP----	EYAKQDF	274	
QY	224	SLMTESYGGHYGPAFFNFHFEQNERIANG	SVNGVQLNFSNLGIINGIIDAIQAPYPEF	283	
Db	275	HIAGESVAGHYIVFASEILSHKKR-----	NINLQSVLNGTUDGYQTYEYRPM	325	
QY	284	AVNNTYGIKAVNETVYNYMKFANQMPNG	QODLICTCKQTNRTALADYALCAEATNMC	343	
Db	326	ACDGGGPVAVLDSESSCQSD--NALPR--	CQSMTESCYSESA-----WVCVPASII	377	
QY	344	VEGPIYAFAGRGVYDIRHPYDOP----	TPPSYNNKFLAKSDVMAIGWIN--YTOS	397	
Db	378	LLAPIOR--TGQNVYDVRGKCEOSSN	LCYAGMVSDYLNKPEVIEAVGAEWNGYD	436	
QY	398	VYVAFQOTGDFVWP--NFIEDLEEILALE	VRVSLIY--GDADYICNFWGGAVALAANS	454	
Db	437	INRNELFHGDWMMKPVHRLVPGLL--	QIPV--LIYAGDADFICNLGNKAWTEALE	491	
QY	455	AAQFERSAGYTPLKV-----NGVEY	GETRYGNFSTRVYEAGHEVYYPQIPASLO	509	

RESULT 4

RESULT 4
US-08-608-267-2

; Sequence 2, Application US/08608267

; Patent No. 5688663

; GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue

APPLICANT: Thompson, Sheryl Ann

284 AVNNTYGIKAVNETVYNYMKFANQMNGCODLLSTCKQTNRTADALVALCAEATNMCRDN 343
 326 ACDDGGYPVAVLDESSQSDM--NALER-COSMIESCYSSSA-----WVCVPASIIYCNNA 377
 344 VEGPYAFAGRGVYDIRHPYDDP-----TPPSYNNKELAKDSVMDAIGVNNIN-YTQSNND 397
 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCYSANGVYSDYLNKPEVIEAAGVNGYDSCNFD 436
 398 VYFAFOOTGDFVWP--NFIEDLEILALPVRVSLIY-GDADYICNWFGGQAVSLAANYSQ 454
 437 INRNFLPHGDMKPYHLVPGLLF--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491
 455 AAQFRSAGYPLKV-----NGVEYGETREYGNRSFTRVYBAGHEVPPYQPIASLQLFNRT 509
 492 QAEYASAELEDLVVDNEHTGKIGQVKSHGNFTFMRLYGGCHVMPMDQPESSLEFFNRW 551
 510 IFG-W 513
 552 LGGEW 556

RESULT 6
 US-08-608-224-2
 ; Sequence 2, Application US/08608224
 ; Patent No. 5705376
 ; GENERAL INFORMATION:
 ; APPLICANT: Yaver, Debbie Sue
 ; APPLICANT: Thompson, Sheryl Ann
 ; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.
 ; STREET: 405 Lexington Avenue, Suite 6400
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/608,224
 ; FILING DATE: 28-FEB-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/309,341
 ; FILING DATE: 20-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney, Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4247.000-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 867 0123
 ; TELEFAX: 212 867 0298
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 557 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Aspergillus Niger
 ; US-08-608-224-2

Query Match 21.1%; Score 630; DB 1; Length 557;
 Best Local Similarity 31.8%; Pred. No. 7.2e-49;
 Matches 154; Conservative 81; Mismatches 174; Indels 76; Gaps 17;

45 ANNVITRYKEFGAGVCCTTGPVKYSYGYVDTSP-ESHTFFWFPEARHNDETAITLWLN 103
 132 AYDLRVKKTDPGSLGI---DFGVKQYTGILDDNENDKHLFWFESENDEPNPVLWLN 198
 104 GPGGSDSLIGLPELGFCHVNSTFDDYINPHSNVSNLFLSQPLGVGFSYSDTVDGSI 163
 189 GPGCCSLTGLFELGFPSSINKTKTPQVNDYAMNSAVIFLDQPVNVGYSYNSA---- 244
 164 NPVTGVVNSSFAGVQRYFTIDATLIDTNLAARAAWEILQGLSLGSLDSRVQSKDF 223
 245 -----VSDTVAAGKVYALTLTFPKQFP-----BYAKQDF 274
 224 SLWTESYGGHYGPAFFNHFEQNERIANGSVNGVQLNFNSLGIINGIDEAIAQAPYYPEF 283
 275 HIAGESYAGHYIPVFASEILLSKKR-----NINLQSVLIGNLTDGYQYVYRPM 325

492 QAEYASAELEDLVVDNEHTGKIGQVKSHGNFTFMRLYGGCHVMPMDQPESSLEFFNRW 551
 510 IFG-W 513
 552 LGGEW 556

RESULT 5
 US-08-608-452-2
 ; Sequence 2, Application US/08608452
 ; Patent No. 5693510
 ; GENERAL INFORMATION:
 ; APPLICANT: Yaver, Debbie Sue
 ; APPLICANT: Thompson, Sheryl Ann
 ; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: No. 56935100 No. 5693510disk of No. 5693510th America, Inc.
 ; STREET: 405 Lexington Avenue, Suite 6400
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/608,452
 ; FILING DATE: 28-FEB-1996
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/309,341
 ; FILING DATE: 20-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney, Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4247.000-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 867 0123
 ; TELEFAX: 212 867 0298
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 557 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Aspergillus Niger
 ; US-08-608-452-2

Query Match 21.1%; Score 630; DB 1; Length 557;
 Best Local Similarity 31.8%; Pred. No. 7.2e-49;
 Matches 154; Conservative 81; Mismatches 174; Indels 76; Gaps 17;

45 ANNVITRYKEFGAGVCCTTGPVKYSYGYVDTSP-ESHTFFWFPEARHNDETAITLWLN 103
 132 AYDLRVKKTDPGSLGI---DFGVKQYTGILDDNENDKHLFWFESENDEPNPVLWLN 198
 104 GPGGSDSLIGLPELGFCHVNSTFDDYINPHSNVSNLFLSQPLGVGFSYSDTVDGSI 163
 189 GPGCCSLTGLFELGFPSSINKTKTPQVNDYAMNSAVIFLDQPVNVGYSYNSA---- 244
 164 NPVTGVVNSSFAGVQRYFTIDATLIDTNLAARAAWEILQGLSLGSLDSRVQSKDF 223
 245 -----VSDTVAAGKVYALTLTFPKQFP-----BYAKQDF 274
 224 SLWTESYGGHYGPAFFNHFEQNERIANGSVNGVQLNFNSLGIINGIDEAIAQAPYYPEF 283
 275 HIAGESYAGHYIPVFASEILLSKKR-----NINLQSVLIGNLTDGYQYVYRPM 325

ADDRESSEE: No. 568663c No. 568663disk of No. 568663th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/608,267
FILING DATE: 28-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger
US-08-608-267-4

Query Match 20.9%; Score 623; DB 1; Length 557;
Best Local Similarity 32.0%; Pred. No. 3.1e-48;
Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

QY 45 ANNVTIRKEPGAEGVCTTPGVKSYGVVDTS-ESHTFFWFFPEARHNPETAPITLWLN 103
Db 132 AYDLVKKTKDPSLSGI--DPGVKQYTGYYLDNDKHLFYWFFSRNDPNDPVVWLN 188
QY 104 GPGSDSLGLFEEELGCHVNSTFDYINPHSWNEVSNLLFSLQPLGVGFSYSDTVDSGI 163
Db 189 GPGGCSLTGLFMELGPFSSINKKIQPVNDYAWNSASVIFLDPQVNVGYSYNSA---- 244
QY 164 NPVTGVVENSFAGVQGRYPTIDATLIDTNNLAEEAAWEILOGLFSLPSLSRSVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274
QY 224 SLWTSYGGHYGPAFFNHFYEQNERJANGSVQVNFNSLGIINGIIDEAIAQAPYFEP 283
Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLSVLIGNLTDGLTQYFYRPM 325
QY 284 AVNNTYGIKAVNETVYNYMKFANQMPGCGDLSTCKQNTRTALADYALCAEATNCRDN 343
Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESCVSESA-----WVCVPASIYCNNA 377
QY 344 VEGPYVAFAGRGVYDIRHPYDDP-----TPPSYNNKFLAKDSYMDAIGVNNIN-YTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCYSAMGVSDYLNKTEVIEAVGAEVNGYDSCNFD 436
QY 398 VYVAFQGTGDFVMP--NFIEDLEILALPVRVSLIY-GDADYICNWFEGGQAVSLAANSYQ 454
Db 437 INRNFLPHGDMKPYHRLVPGLE--QIPV--LIVAGDADFICNWLGNKAWTEALEWPG 491
QY 455 AAQFRSAGYTPLVK-----NGVEYGETREYGNFSFTRVYBAGHEVPYQYIASQLFNRT 509
Db 492 QAEYASAKLEDLVVNEHKGKIGQVKSNGHNTFMELYGCGHVPMDQPESSLEFFNRW 551
QY 510 IFG-W 513
552 LGGEW 556

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/309,341
FILING DATE: 16-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger
US-08-309-341-4

Query Match 20.9%; Score 623; DB 1; Length 557;
Best Local Similarity 32.0%; Pred. No. 3.1e-48;
Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

QY 45 ANNVTIRKEPGAEGVCTTPGVKSYGVVDTS-ESHTFFWFFPEARHNPETAPITLWLN 103
Db 132 AYDLVKKTKDPSLSGI--DPGVKQYTGYYLDNDKHLFYWFFSRNDPNDPVVWLN 188
QY 104 GPGSDSLGLFEEELGCHVNSTFDYINPHSWNEVSNLLFSLQPLGVGFSYSDTVDSGI 163
Db 189 GPGGCSLTGLFMELGPFSSINKKIQPVNDYAWNSASVIFLDPQVNVGYSYNSA---- 244
QY 164 NPVTGVVENSFAGVQGRYPTIDATLIDTNNLAEEAAWEILOGLFSLPSLSRSVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274
QY 224 SLWTSYGGHYGPAFFNHFYEQNERJANGSVQVNFNSLGIINGIIDEAIAQAPYFEP 283
Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLSVLIGNLTDGLTQYFYRPM 325
QY 284 AVNNTYGIKAVNETVYNYMKFANQMPGCGDLSTCKQNTRTALADYALCAEATNCRDN 343
Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESCVSESA-----WVCVPASIYCNNA 377
QY 344 VEGPYVAFAGRGVYDIRHPYDDP-----TPPSYNNKFLAKDSYMDAIGVNNIN-YTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCYSAMGVSDYLNKTEVIEAVGAEVNGYDSCNFD 436
QY 398 VYVAFQGTGDFVMP--NFIEDLEILALPVRVSLIY-GDADYICNWFEGGQAVSLAANSYQ 454
Db 437 INRNFLPHGDMKPYHRLVPGLE--QIPV--LIVAGDADFICNWLGNKAWTEALEWPG 491
QY 455 AAQFRSAGYTPLVK-----NGVEYGETREYGNFSFTRVYBAGHEVPYQYIASQLFNRT 509
Db 492 QAEYASAKLEDLVVNEHKGKIGQVKSNGHNTFMELYGCGHVPMDQPESSLEFFNRW 551
QY 510 IFG-W 513
552 LGGEW 556

RESULT 9
US-08-608-267-4
Sequence 4, Application US/08608267
Patent No. 568663
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

Db 552 LGGEW 556

RESULT 10

US-08-608-452-4

Sequence 4, Application US/08608452

Patent No. 5693510

GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue

APPLICANT: Thompson, Sheryl Ann

TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 56935100 No. 5693510disk of No. 5693510th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/608,452

FILING DATE: 28-FEB-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/309,341

FILING DATE: 20-SEP-1994

NAME: Lowney, Karen A.

REGISTRATION NUMBER: 31,274

REFERENCE/DOCKET NUMBER: 4247.000-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORGANISM: Aspergillus Niger

US-08-608-452-4

Query Match

Best Local Similarity 20.9%; Score 623; DB 1; Length 557;

Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

QY 45 ANNVTIRKPEAGVCGCTTGGKSGYGVDTSP-ESHTFFWFPEARNHPETATLWLN 103

Db 132 AYDLRVKKTDPSSLGI---DPGVKQYTGYYLDNDNKHLYFWFFSRNDPNDPVVLN 188

QY 104 GPGGSDSLIGLFEELGPGCHVNSTEDDYINPHSWNEVSNLLFLSQPLGVGFYSYSDTVDSGI 163

Db 189 GPGGSSLTGLFMELGPGSINKKIOPVYNDVWNSASVIFLDQPVNVGYYSNSA----244

QY 164 NPVTGVNENSSFAGVQGRYPTIDATLIDTTLAAEAWEILQGLPSLDLSRVSQKDF 223

Db 245 -----VSDTVAAGKDVALLTLFPKQP-----EVAKQDF 274

QY 224 SLWTESYGGHYPFAFFNPFYQNERIANGSVNGVOLFNSLGIINGIIDEAIQAPYYPEF 283

Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLQSVLIGNGLDGLTQVEYRPM 325

QY 284 AVNNTYGIKAVNETVYVYKFAQNPNGCODLSTCKQKTNTALADYALCAEATMCRDN 343

Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESYSSESA-----WVCVPASTYCNNA 377

QY 344 VEGPYAFAGRGVYDIRHPYDDP-----TPPSYNNKELAKDSVMDAIGVNIN-YTQSNND 397

Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCSAMGYVSDYLNKTEVIEAGVNGYDSCNFD 436

QY 398 VYAFQOTGDFVMP--NFIEDLEILALPVRVSLIY-GDADYICNWFPGGQAVSLAANSQ 454

Db 437 INRNFLPHGDMKPYHRLVFLLE--QIPV---LIYAGDADFICNMLGNKAWTALEWPG 491

QY 455 AAQFRSAGYPLKV-----NGVYGETREYGNFSFTRVVEAGHEVPYQYPIASLOLEPRT 509

Db 492 QAEVASAKLEDLVVENEHKGKIGQVKSNGNFTFMRLYCGGHVPMQDPSSLEFPNRW 551

QY 510 IFG-W 513

Db 552 LGGEW 556

RESULT 11

US-08-608-224-4

Sequence 4, Application US/08608224

Patent No. 5705376

GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue

APPLICANT: Thompson, Sheryl Ann

TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/608,224

FILING DATE: 28-FEB-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/309,341

FILING DATE: 20-SEP-1994

NAME: Lowney, Karen A.

REGISTRATION NUMBER: 31,274

REFERENCE/DOCKET NUMBER: 4247.000-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORGANISM: Aspergillus Niger

US-08-608-224-4

Query Match

Best Local Similarity 20.9%; Score 623; DB 1; Length 557;

Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

QY 45 ANNVTIRKPEAGVCGCTTGGKSGYGVDTSP-ESHTFFWFPEARNHPETATLWLN 103

Db 132 AYDLRVKKTDPSSLGI---DPGVKQYTGYYLDNDNKHLYFWFFSRNDPNDPVVLN 188

QY 104 GPGGSDSLIGLFEELGPGCHVNSTEDDYINPHSWNEVSNLLFLSQPLGVGFYSYSDTVDSGI 163

Db 189 GPGGSSLTGLFMELGPGSINKKIOPVYNDVWNSASVIFLDQPVNVGYYSNSA----244

QY 164 NPVTGVNENSSFAGVQGRYPTIDATLIDTTLAAEAWEILQGLPSLDLSRVSQKDF 223

Db 245 -----VSDTVAAGKDVALLTLFPKQP-----EVAKQDF 274

QY 224 SLWTESYGGHYPFAFFNPFYQNERIANGSVNGVOLFNSLGIINGIIDEAIQAPYYPEF 283

Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLQSVLIGNGLDGLTQVEYRPM 325

QY 284 AVNNTYGIKAVNETVYVYKFAQNPNGCODLSTCKQKTNTALADYALCAEATMCRDN 343

Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESYSSESA-----WVCVPASTYCNNA 377

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger
US-08-967-149-4

Query Match 20.9%; Score 623; DB 2; Length 557;
Best Local Similarity 32.0%; Pred. No. 3.1e-48;
Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

Qy 45 ANNVITYIKKPGAGVCETTPGVKSYSGYVDTSP-ESEHTEFWPFEARHNPETAPITLWLN 103
Db 132 AYDLURVKTPFPSSLGI---DPGVKQYTYGLDDNNDKHLFWFESRNDPNDPVLWLN 188
Qy 104 GPGCSLSLIGLFEELGCHVNSTPDDYINPHSMNEVSNLLFSLQPLGVGFYSYDTPDGI 163
Db 189 GPGCSSLTGLFMELGFSSINKKIQPVYNDVAMNSASVIFLDQPVNVGYSYNSA---- 244
Qy 164 NPVTGVVNSSFAGVQGRYPTIDATLDTTNLAEEAAWEILQGLSLGLPSLDSRVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274
Qy 224 SLWTESYGGHYGPAFFNFHFEQNERIANGSVNGVQLNENSLGIINGIIDEAIQAPYPEF 283
Db 275 HIAGESYAGHYIPVFAEILSHKXR-----NINLQSVLIGNGLTDGLTQYEYRPM 325
Qy 284 AVNNTYIGKAVNETVYNYMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNCRDN 343
Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQMIESCYSSSA-----WVCVPASTYCNA 377
Qy 344 VEGPIYAFAGRGVDIRHPYDDP-----TPSYNKNFLAKOSVMDAIGVNIIN-VTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCSAMGYVSDYLNKTEVIEAAGAEVNGYDSCNFD 436
Qy 398 VYVAFQQTGDFWMP-NFIEDLEETILALPVRVSLIY-GDADYICNWFQGGQAVSLAANYSQ 454
Db 437 INRNFLFGDNKPKYHRLVPGLE--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491
Qy 455 AAQFRSAGYTPLKV-----NGVEYGETREYNFSTRVYEAGHEVPYIOPIASLQLFNRT 509
Db 492 QAEYASAKLEDLVVENEHKGKIGQVKSHGNFTFMRLYGGGHVMPMDQPSSELEFFNRW 551
Qy 510 IFG-W 513
Db 552 LGGEW 556

RESULT 13
US-09-640-305-4
Sequence 4, Application US/09640305
Patent No. RE37447
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Yeh, Patrice
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
PREPARATION AND USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

189 GPGCSSLTGLFMELGFSSINKKIQPVYNDVAMNSASVIFLDQPVNVGYSYNSA---- 244
Qy 164 NPVTGVVNSSFAGVQGRYPTIDATLDTTNLAEEAAWEILQGLSLGLPSLDSRVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274
Qy 224 SLWTESYGGHYGPAFFNFHFEQNERIANGSVNGVQLNENSLGIINGIIDEAIQAPYPEF 283
Db 275 HIAGESYAGHYIPVFAEILSHKXR-----NINLQSVLIGNGLTDGLTQYEYRPM 325
Qy 284 AVNNTYIGKAVNETVYNYMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNCRDN 343
Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQMIESCYSSSA-----WVCVPASTYCNA 377
Qy 344 VEGPIYAFAGRGVDIRHPYDDP-----TPSYNKNFLAKOSVMDAIGVNIIN-VTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCSAMGYVSDYLNKTEVIEAAGAEVNGYDSCNFD 436
Qy 398 VYVAFQQTGDFWMP-NFIEDLEETILALPVRVSLIY-GDADYICNWFQGGQAVSLAANYSQ 454
Db 437 INRNFLFGDNKPKYHRLVPGLE--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491
Qy 455 AAQFRSAGYTPLKV-----NGVEYGETREYNFSTRVYEAGHEVPYIOPIASLQLFNRT 509
Db 492 QAEYASAKLEDLVVENEHKGKIGQVKSHGNFTFMRLYGGGHVMPMDQPSSELEFFNRW 551
Qy 510 IFG-W 513
Db 552 LGGEW 556

RESULT 12
US-08-967-149-4
Sequence 4, Application US/08967149
Patent No. 5939305
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5939305disk of No. 5939305th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,149
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,452
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowrey, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids

RESULT 14
US-08-360-673-4
; Sequence 4, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:51:55 ; Search time 46 Seconds
 (without alignments)
 1915.071 Million cell updates/sec

Title: US-09-712-338-2
 Perfect score: 2979
 Sequence: 1 MRGVEFLSVLPVLAASWALP.....HTQSSVPLPATSMSVGM 555

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :

A_Geneseq 19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2975	99.9	554	19	AAW56099
2	2376	79.8	551	23	ABR38864
3	980.5	32.9	523	23	ABR38865
4	941	31.6	554	23	ABR38859
5	925	31.1	526	23	ABR38817
6	636	21.3	508	23	ABG93049
7	630	21.1	536	23	ABR38819
8	630	21.1	556	17	AA96737
9	620	20.8	557	17	AA96738
10	554	19	AAW56099		A. oryzae ATCC2038
11	551	23	ABR38864		A. niger serine ca
12	523	23	ABR38865		A. niger serine ca
13	554	23	ABR38859		A. niger serine ca
14	526	23	ABR38817		A. niger serine ca
15	508	23	ABG93049		S. cerevisiae BAX-
16	536	23	ABR38819		A. niger carboxype
17	556	17	AA96737		A. niger Bo-1 carb
18	557	17	AA96738		A. niger SFAG 2 ca

10	568.5	19.1	491	15	AA48059	Sequence of protea
11	541.5	18.2	550	23	ABG93281	C. albicans BAX-as
12	419	14.1	481	23	ABR38845	A. niger carboxype
13	383	12.9	482	21	AAG30065	Arabidopsis thalia
14	383	12.9	502	21	AAG30064	Arabidopsis thalia
15	365	12.3	455	23	ABR38843	A. niger carboxype
16	361.5	12.1	479	21	AAG31905	Arabidopsis thalia
17	361.5	12.1	486	21	AAG31904	Arabidopsis thalia
18	361	12.1	476	19	AAW72966	Human serine carbo
19	361	12.1	476	22	AA88381	Human membrane or
20	359	12.1	476	20	AA133372	Amino acid sequenc
21	359	12.1	476	21	AA801407	Human TANGO 176.
22	359	12.1	476	22	AAU29228	Human PRO polypept
23	359	12.1	476	22	AA88587	Human hydrophobic
24	359	12.1	476	22	AA80240	Human PRO223 prote
25	359	12.1	476	24	ABU69650	Novel human secret
26	359	12.1	476	24	ABU71316	Human PRO polypept
27	359	12.1	476	24	ABU71473	Human secreted/tra
28	359	12.1	476	24	ABU71919	Human secreted/tra
29	359	12.1	476	24	ABU65773	Novel human secret
30	359	12.1	476	24	ABU66106	Human secreted/tra
31	359	12.1	476	24	ABU67373	Human secreted/tra
32	359	12.1	476	24	ABU67610	Human secreted pro
33	359	12.1	476	24	ABU64527	Human secreted/tra
34	359	12.1	476	24	ABU65468	Human secreted/tra
35	359	12.1	476	24	ABU58604	Human PRO polypept
36	359	12.1	476	24	ABU56140	Human secreted/tra
37	359	12.1	476	24	ABU57135	Human PRO polypept
38	359	12.1	476	24	ABU54375	Human secreted/tra
39	359	12.1	476	24	ABU10714	Human secreted/tra
40	359	12.1	477	22	AAW25810	Human protein sequ
41	357.5	12.0	479	21	AA474778	Arabidopsis thalia
42	357.5	12.0	486	21	AA474777	Arabidopsis thalia
43	357	12.0	476	20	AAU28570	Secreted peptide c
44	357	12.0	476	22	AAU39043	Human secreted pro
45	357	12.0	476	23	AB555752	Human polypeptide

ALIGNMENTS

RESULT 1

AAW56099
 ID AAW56099 standard; Protein; 554 AA.

XX AAW56099;
 AC AAW56099;

XX 27-AUG-1998 (first entry)
 DT

XX A. oryzae ATCC20386 carboxypeptidase I protein.
 DE

XX Carboxypeptidase I; flavour improving agent; hydrolysat; proteinaceous;
 KW food industry.
 XX

OS Aspergillus oryzae.
 XX

XX Key Location/Qualifiers
 FH Peptide 1..18
 FT /label= signal
 FT Protein 19..554
 FT /label= carboxypeptidase I

XX WO9814599-A1.
 PN

XX 09-APR-1998.
 PD

XX 03-OCT-1997; 97WO-US17977.
 PF

XX 27-NOV-1996; 96US-0757534.
 PR

XX 04-OCT-1996; 96US-0726880.
 PR

XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA

XX (NOVO) NOVO-NORDISK AS.
 PA

[illegible]

WPI; 2002-723203/78.

AC	XX	ABR38859;	
XX	XX		
DI	XX	24-APR-2003 (first entry)	
DE	XX	A. niger serine carboxypeptidase polypeptide #2.	
XX	XX	Protease; fungal infection; aspergillosis; food; tanning; detergent;	
XX	XX	protein solubility; viscosity; taste; texture; nutritional value;	
KW	KW	EC3.4.16.6.	
KW	XX		
OS	XX	Aspergillus niger.	
XX	XX		
PN	XX	WO200268623-A2.	
XX	XX		
PD	XX	06-SEP-2002.	
XX	XX		
PD	XX	22-FEB-2002; 2002WO-EP01984.	
XX	XX		
XX	PR	23-FEB-2001; 2001EP-0200657.	
XX	PR	23-FEB-2001; 2001EP-0200658.	
XX	PR	23-FEB-2001; 2001EP-0200660.	
XX	PR	26-FEB-2001; 2001EP-0200706.	
XX	PR	26-FEB-2001; 2001EP-0200707.	
XX	PR	26-FEB-2001; 2001EP-0200708.	
XX	PR	26-FEB-2001; 2001EP-0200719.	
XX	PR	28-MAR-2001; 2001EP-0000075.	
XX	PR	28-MAR-2001; 2001EP-0000078.	
XX	PR	28-MAR-2001; 2001EP-0000080.	
XX	PR	28-MAR-2001; 2001EP-0000087.	
XX	PR	28-MAR-2001; 2001EP-0000088.	
XX	PR	21-MAY-2001; 2001EP-0000156.	
XX	PR	21-MAY-2001; 2001EP-0000159.	
XX	PR	21-MAY-2001; 2001EP-0000160.	
XX	PR	21-MAY-2001; 2001EP-0000162.	
XX	PR	21-MAY-2001; 2001EP-0000165.	
XX	PR	21-MAY-2001; 2001EP-0000166.	
XX	PR	21-MAY-2001; 2001EP-0000168.	
XX	PR	21-JUN-2001; 2001EP-0000240.	
XX	PR	21-JUN-2001; 2001EP-0000242.	
XX	PR	21-JUN-2001; 2001EP-0000244.	
XX	PR	21-JUN-2001; 2001EP-0000246.	
XX	PR	12-JUL-2001; 2001EP-0000280.	
XX	PR	12-JUL-2001; 2001EP-0000285.	
XX	PR	30-JUL-2001; 2001EP-0000323.	
XX	PR	30-JUL-2001; 2001EP-0000327.	
XX	PR	02-AUG-2001; 2001EP-0000341.	
XX	PR	02-AUG-2001; 2001EP-0000342.	
XX	PR	02-AUG-2001; 2001EP-0000343.	
XX	PR	02-AUG-2001; 2001EP-0000344.	
XX	PR	09-AUG-2001; 2001EP-0000357.	
XX	PR	16-AUG-2001; 2001EP-0000374.	
XX	PR	16-AUG-2001; 2001EP-0000377.	
XX	PR	20-SEP-2001; 2001EP-0000478.	
XX	PR	20-SEP-2001; 2001EP-0000483.	
XX	PR	22-OCT-2001; 2001EP-0000552.	
XX	PR	22-OCT-2001; 2001EP-0000553.	
XX	PR	22-OCT-2001; 2001EP-0000554.	
XX	PR	22-OCT-2001; 2001EP-0000556.	
XX	PR	22-OCT-2001; 2001EP-0000557.	
XX	PR	22-OCT-2001; 2001EP-0000558.	
XX	PR	15-NOV-2001; 2001EP-0000464.	
XX	PR	21-DEC-2001; 2001EP-0005117.	
XX	XX	(STAM) DSM NV.	
XX	XX		
XX	PI	Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel	
XX	PI	Klugbauer S, Wagner C, Fritz A, von Gustedt W, Heinrich O;	
XX	PI	Maler D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;	
XX	PI	Stiebler J, Albang R;	
XX	XX		
DR	DR	WPI; 2002-723203/78.	
DR	DR	P-PSDB; ABZ78226, ABZ78283.	

PT Novel isolated protease polypeptide useful in laboratory, clinical,
 PT pharmaceutical, chemical, diagnostic, personal care and industrial
 XX applications -
 XX
 PS Claim 13; Page 366-369; 394pp; English.
 CC
 CC The invention relates to a novel isolated protease polypeptide. A
 CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 CC fungal infection such as aspergillosis, or as a query sequence to perform
 CC a search against public databases. A polypeptide of the invention is
 CC useful in a selected number of industrial or pharmaceutical processes, in
 CC laboratory or clinical processes, in food industry (baking, brewing,
 CC cheese manufacturing, meat tenderising), in tanning industry and in the
 CC manufacture of biological detergents. A polypeptide may also be useful
 CC for improving protein solubility, extraction yields, viscosity or taste,
 CC texture, nutritional value, minimising of antigenicity or
 CC anti-nutritional factors, colour or functionality as well as processing
 CC aspects like filterability of the proteinaceous raw material. The
 CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
 CC the invention.
 XX
 SQ Sequence 554 AA;
 Query Match 31.6%; Score 941; DB 23; Length 554;
 Best Local Similarity 38.5%; Pred. No. 1.2e-73;
 Matches 209; Conservative 72; Mismatches 172; Indels 90; Gaps 13;
 QY 31 QLPKNTGVKLTFA--NNVTIRKPGCAEVCCTTPGVKSYGVYDTSPEHFFWFFE 88
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 88
 17 QFPEPEGITVLKSLHENVITISKEP--GICETTFGVRSYGVYHLPASTSFFWFFE 73
 QY 89 ARHNPEAPITLWNGPGSDSLGLFEELGPCHVNS--TFDDYINPHSMNEVSNLLFSLQ 147
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 147
 74 ARKDPNAPLAIWNGPGSSSLMLLEELGPCSIASDSKTTVLNPSWNEVSNLLFLDQ 133
 QY 148 PLGVGFSYSTVDGSIINPVTGVVNSSFAGVQGYPTIDAILDITNLAEEAWEILQGF 207
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 207
 134 PTQVGFSDYPTNGTLT-----ANGTAFAAHALWHEAQTW 168
 QY 208 LSGLSL---DSRVQSKDFSLWTSYGGHYGPAFFNFHYEONERIANGSV-NGVQ-LNFN 262
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 262
 169 FFEFFHYKPNDRV-----SLWAESGGHYGPGIFRFFQOQNDKIAEGTAEDGAQYLHLD 223
 QY 263 SLGIINGIDEALQAPYPEFVANNVTYGIKAVNETVYNYMKFANOMPNGCQDLISCTKQT 322
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 322
 224 TLGIVNGLMDVIOEEAYITWP-----YNNVRLAP-----SSF 256
 QY 323 NRTALADYALCAEATNCRD-----NV-----EGP--YFAPAGRGVYDIR 360
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360
 257 NSRGFRDQALACEALKERDSGLPHSGKNISEICGGLALEWGDGPITYYHTFNRGWYDIA 316
 QY 361 HPYDDPTPSYNNKFLAKDSVMDAIGWINYNTOSNDVYVAFQOCTGDFVKNPFIEDLEI 420
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 420
 317 HPKNOPFPKAKMLGYLTQESVLAALGVPNVTSSSAVATQFTKTDIVHGFLDAIGYL 376
 QY 421 LALPVRSVLIYGDYICNWFQGVQAVSLAANYSQAAQFRSAGYTPILKVNVEYGETREYG 480
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 480
 377 LDGSKVHMVYGDYACNVWGEKASLAVPYSRITEFADTGYSPLITPDGISGMRQLG 436
 QY 481 NFSFTRVVEAGHEVPYQYQIASLQFNRTIFGWDIAEGQKKIWPYSKYNGTATATHTQSS 540
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 540
 437 NYSFTRVFOAGHEVFSYQPVAAIEFMRAFKNKDIPTGLLAVDDBFQSVGPKDTHIKNI 496
 QY 541 VPL 543
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 496
 497 PPI 499
 RESULT 5
 ABR38817
 ID ABR38817 standard; Protein; 526 AA.
 XX
 AC ABR38817;
 PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebler J, Albang R;
 XX WPI; 2002-723203/78.
 DR P-PSDB; ABZ78184, ABZ78241.
 XX Novel isolated protease polypeptide useful in laboratory, clinical,
 PT

XX
 DT 24-APR-2003 (first entry)
 XX
 DE A. niger serine carboxypeptidase polypeptide #1.
 XX
 KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.6.
 XX
 OS Aspergillus niger.
 XX
 FN WO200268623-A2.
 XX
 PD 06-SEP-2002.
 XX
 XX 22-FEB-2002; 2002WO-EP01984.
 PF 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0000075.
 PR 28-MAR-2001; 2001EP-0000078.
 PR 28-MAR-2001; 2001EP-0000080.
 PR 28-MAR-2001; 2001EP-0000087.
 PR 21-MAY-2001; 2001EP-0000088.
 PR 21-MAY-2001; 2001EP-0000156.
 PR 21-MAY-2001; 2001EP-0000159.
 PR 21-MAY-2001; 2001EP-0000160.
 PR 21-MAY-2001; 2001EP-0000162.
 PR 21-MAY-2001; 2001EP-0000165.
 PR 21-MAY-2001; 2001EP-0000166.
 PR 21-MAY-2001; 2001EP-0000168.
 PR 21-JUN-2001; 2001EP-0000240.
 PR 21-JUN-2001; 2001EP-0000242.
 PR 21-JUN-2001; 2001EP-0000244.
 PR 21-JUN-2001; 2001EP-0000246.
 PR 12-JUL-2001; 2001EP-0000280.
 PR 12-JUL-2001; 2001EP-0000285.
 PR 30-JUL-2001; 2001EP-0000323.
 PR 30-JUL-2001; 2001EP-0000327.
 PR 02-AUG-2001; 2001EP-0000341.
 PR 02-AUG-2001; 2001EP-0000342.
 PR 02-AUG-2001; 2001EP-0000343.
 PR 02-AUG-2001; 2001EP-0000344.
 PR 09-AUG-2001; 2001EP-0000357.
 PR 16-AUG-2001; 2001EP-0000374.
 PR 16-AUG-2001; 2001EP-0000377.
 PR 20-SEP-2001; 2001EP-0000478.
 PR 20-SEP-2001; 2001EP-0000483.
 PR 22-OCT-2001; 2001EP-0000552.
 PR 22-OCT-2001; 2001EP-0000553.
 PR 22-OCT-2001; 2001EP-0000554.
 PR 22-OCT-2001; 2001EP-0000556.
 PR 22-OCT-2001; 2001EP-0000557.
 PR 22-OCT-2001; 2001EP-0000558.
 PR 15-NOV-2001; 2001EP-0004464.
 PR 21-DEC-2001; 2001EP-0005117.
 XX (STAM) DSM NV.
 PA Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 XX Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebler J, Albang R;
 XX WPI; 2002-723203/78.
 DR P-PSDB; ABZ78184, ABZ78241.
 XX Novel isolated protease polypeptide useful in laboratory, clinical,
 PT

QY	284	AVNNTYGIKAVNETVYMKFANOMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDN	343
Db	305	ACGGGYPVAVLDESSCQMD--NALPR-CQSMIESCYSESA-----WVCVPASYICNNA	356
QY	344	VEGPIYAFAGRGVDIRHEYPDP-----TPPSYNNKFLAKSVMDAIGVNNIN-YTQSNND	397
Db	357	LLAPYQR-TGQNVYDVRGKCEDSSNLCYSAMGYSDYLNKPEVIEAVGAENVGYDSCNPD	415
QY	398	VYIAFOQTGDFWVP--NFIEDLEELALPVRVSLIY-GDADYICNWFEGGQAVSLAANYSQ	454
Db	416	INRNFLFHGDWMPYHRLVPGLE--QIPV---LIYAGDADFCINWLGKAWTEALEWPG	470
QY	455	AAQPSAGVTPLKV-----NGVEYGETREYGNFSTFRVYEAGHEVPYQPIASLQLFNRT	509
Db	471	QAEYASAELEDLVIVDNEHTGKKIQVKGSHGNFTFMRLYGGGHVMPMDQPESSLEFFNRW	530
QY	510	IFG-W 513	
Db	531	LGGEW 535	
RESULT 8			
ID	AAR96737	standard; Protein; 556 AA.	
AC	AAR96737;		
XX	14-AUG-1996	(first entry)	
XX	A. niger	Bo-1 carboxypeptidase Y.	
XX	Carboxypeptidase Y; CPY; ascomycete; deuteromycete; host cell;		
XX	protease deficiency.		
XX	Aspergillus niger	strain Bo-1.	
XX	WO9609397-A1.		
XX	28-MAR-1996.		
XX	19-SEP-1995;	95WO-US11945.	
XX	20-SEP-1994;	94US-0309341.	
XX	(NOVO)	NOVO NORDISK BIOTECH INC.	
XX	Thompson SA,	Yaver DS;	
XX	WPI; 1996-188458/19.		
XX	N-PSDB; AAT28283.		
XX	Nucleic acid construct encoding a filamentous ascomycete or		
XX	deuteromycete carboxy:peptidase Y - useful to produce host cells		
XX	modified to produce reduced amounts of carboxy:peptidase		
XX	Claim 4; Fig 1A-E; 46pp; English.		
XX	Carboxypeptidase Y (AAR96737), a vacuolar protease, is the product of		
XX	the CPY gene (AAT28283) of Aspergillus niger strain Bo-1. Creation		
XX	of CPY-deficient Aspergillus strains, e.g. by cloning a selectable		
XX	marker into the CPY gene, provides suitable host strains for prodn.		
XX	of heterologous protein.		
XX	Sequence	556 AA;	
XX	Query Match	21.1%; Score 630; DB 17; Length 556;	
XX	Best Local Similarity	31.8%; Pred. No. 2.5e-46;	
XX	Matches 154; Conservative	81; Mismatches 174; Indels 76; Gaps 17;	
QY	45	ANNVTIRYKEPGAEGVCETTPGKSYSGYVDTP--ESHTFFWFFFEARHNPTAPITLWLN	103
Db	131	AYDLRVKKTDPGSLGI---DPGVKQYTGVLDDNENDKHLFYWFFESRNDPNDPVVLWLN	187

QY	104	GGPSDSLIGLFEELGPHCHVNSTFDDYINPHSMWNEVSNLLFLSQPLGVGFYSYSDVDSGI	163
Db	188	GGPGCSSTGLFMELGPSSINKKIQPVNDYANNNSASVIFLDQPVNVGVYSYNSA----	243
QY	164	NPVTGVWVENSFAGVQGRYPTIDATLIDTNNLAAPAAWEILOGLFSLPSLDSRVQSKDF	223
Db	244	-----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF	273
QY	224	SLMTSEYGHVGPAPFNHFYQNERLANGSVNGVQLNENSLGIINGLIIDEALQAPYYPEF	283
Db	274	HIAGESYAGHYIPVFASBILSHKKR-----NINLQSVLIGNGLTDGTYQYIYRPM	324
QY	284	AVNNTYGIKAVNETVYMKFANOMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDN	343
Db	325	ACGGGYPVAVLDESSCQMD--NALPR-CQSMIESCYSESA-----WVCVPASYICNNA	376
QY	344	VEGPIYAFAGRGVDIRHEYPDP-----TPPSYNNKFLAKSVMDAIGVNNIN-YTQSNND	397
Db	377	LLAPYQR-TGQNVYDVRGKCEDSSNLCYSAMGYSDYLNKPEVIEAVGAENVGYDSCNPD	435
QY	398	VYIAFOQTGDFWVP--NFIEDLEELALPVRVSLIY-GDADYICNWFEGGQAVSLAANYSQ	454
Db	436	INRNFLFHGDWMPYHRLVPGLE--QIPV---LIYAGDADFCINWLGKAWTEALEWPG	490
QY	455	AAQPSAGVTPLKV-----NGVEYGETREYGNFSTFRVYEAGHEVPYQPIASLQLFNRT	509
Db	491	QAEYASAELEDLVIVDNEHTGKKIQVKGSHGNFTFMRLYGGGHVMPMDQPESSLEFFNRW	550
QY	510	IFG-W 513	
Db	551	LGGEW 555	
RESULT 9			
ID	AAR96738	standard; Protein; 557 AA.	
AC	AAR96738;		
XX	14-AUG-1996	(first entry)	
XX	A. niger	SPAG 2 carboxypeptidase Y.	
XX	Carboxypeptidase Y; CPY; ascomycete; deuteromycete; host cell;		
XX	protease deficiency.		
XX	Aspergillus niger	strain SPAG 2.	
XX	Key	Location/Qualifiers	
XX	Peptide	1..138	
XX	Protein	/label= Pre-pro_peptide	
XX	WO9609397-A1.	139..557	
XX	28-MAR-1996.	/label= Mat_protein	
XX	19-SEP-1995;	95WO-US11945.	
XX	20-SEP-1994;	94US-0309341.	
XX	(NOVO)	NOVO NORDISK BIOTECH INC.	
XX	Thompson SA,	Yaver DS;	
XX	WPI; 1996-188458/19.		
XX	N-PSDB; AAT28284.		
XX	Nucleic acid construct encoding a filamentous ascomycete or		
XX	deuteromycete carboxy:peptidase Y - useful to produce host cells		
XX	modified to produce reduced amounts of carboxy:peptidase		

PS Claim 5; Page 23-25; 46pp; English.

XX Carboxypeptidase Y (AAR96738), a vacuolar protease, is the product of
CC the CPY gene (AAT28284) of *Aspergillus niger* strain SFAG 2. Creation
CC of CPY-deficient *Aspergillus* strains, e.g. by cloning a selectable
CC marker into the CPY gene, provides suitable host strains for prodn.
XX of heterologous protein.

XX SQ Sequence 557 AA;

Query Match 20.8%; Score 620; DB 17; Length 557;
Best Local Similarity 31.8%; Pred. No. 1.9e-45;
Matches 154; Conservative 80; Mismatches 175; Indels 76; Gaps 18;

Qy 45 ANNVTIRYKEPGAEGVCETTPGVKSYGVVDTSPESTFFWFEAREHNPETAPITLWN 103
Db :
132 AYDLRVKKTDPSSIGI---DPGVKYTGVLDDNNDKHLYFWFFSRNDPENDPVVLELN 188
Qy 104 GPGCSLSLGLFELGPFCHVNSTFDYINPHSWNEVSNIILFLSQPLGVGFYSYDVTGSI 163
Db :
189 GPGCSSSLTGLFMELGPSSINKKIQPVNDYAWNNSASVIFLDQPVGVSNSA---- 244
Qy 164 NPVTGVVNSSPAGVQGRPTIDATLIDTTNLAAEAWEILOGLSGLPSLDSRVQSDF 223
Db :
245 -----VSDTVAAKGDVALLTLFFKQFP-----EYAKQDF 274
Qy 224 SLMTESYGHHGYPAFFNFHFVEQNRIANGSVNGVQLNFNSLGIINGIIDALQAPYYEF 283
Db :
275 HIAGESYAGHYIPVFASEILSHKR-----NINQSVLIGNGLTDGLTEYEVYRPM 325
Qy 284 AVNNTYGIKAVNETVYNMKFANQMPCGDLIISTCKQTNRTALADYALCAEATMCRDN 343
Db :
326 ACDGG-GYPAVLDE-GSCQANDNALPR-CQSMIESCYSSSA-----WVCVPASIYCNA 377
Qy 344 VEGPYAFAGRGVYDIRHPYDDP-----TPPSYNNKFLLAKDSVMDAIGVIN--YTQSNND 397
Db :
378 LLAPYOR-TQNVYDVRGKCDSNLCYSAMGYSDYLNKTEVIEAAGVNYGDCNED 436
Qy 398 VYFAFOQTGDFVMP--NFIEDLEEIALPVRVSLIY-GDADYICNWFGQAVSLAANYSQ 454
Db :
437 INRNFLFHGDWKPKYHLRVPGLE--QIPV---LIYAGDAFCINWLGNKAWTBALRWPG 491
Qy 455 AQCFRSAGYTPLKV-----NGVEYGETREYGNFSFRVYEAGHEVPPYQPIASIQLENRT 509
Db :
492 QAQYASAKLEDLVVENEHBKKIKGVKSHGNFTFMRLYGGGHVMPDQPESSLEFFNRW 551
Qy 510 IFG-W 513
Db :
552 LGGEW 556

RESULT 10
AAR48059
ID AAR48059 standard; Protein; 491 AA.
AC AAR48059;
DT 25-MAR-2003 (updated)
XX 20-JUL-1994 (first entry)
Sequence of protease C encoded by gene Kl.PRC1.
DE DE
XX Protease; yeast; proteolysis.
KW Kluyveromyces lactis.
OS Kluyveromyces lactis.
XX W09400579-A1.
PD 06-JAN-1994.
XX 23-JUN-1993; 93WO-FR00623.
PF XX
XX 25-JUN-1992; 92FR-0007785.
PR XX

(RHON) RHONE POULENC RORER SA.
Fleer R, Fournier A, Yeh P;
WPI; 1994-026215/03.
DR N-PSDB; AAQ55347.
XX New Kluyveromyces yeast with modified protease gene - esp. used
PT for high yield prodn. of recombinant protein, also DNA encoding
PT yeast protease and derived peptide(s)
PS Disclosure; Page 28-31; 49pp; English.
XX The protease gene is to be modified in order to render it (partially)
CC incapable of producing the natural protein; or result in a non-
CC functional protease or in a protease with modified proteolytic
CC activity. The modifications can be introduced in vitro or in situ by
CC standard genetic engineering techniques or by exposure to mutagenic
CC agents.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 491 AA;

Query Match 19.1%; Score 568.5; DB 15; Length 491;
Best Local Similarity 31.1%; Pred. No. 5.2e-41;
Matches 151; Conservative 66; Mismatches 184; Indels 85; Gaps 14;

Qy 45 ANNVTIRYKEPGAEGVCETTPGVKSYGVVDTSPESTFFWFEAREHNPETAPITLWN 104
Db :
62 AYSLRKLPDKSLGV---DTVKOWSGYLDYQSKHFFWFERSNDPENDPVILMLNG 117
Qy 105 GPGCSLSLGLFELGPFCHVNSTFDYINPHSWNEVSNIILFLSQPLGVGFYSYDVTGSI 164
Db :
118 GPGCSSFVGLFELGPFSSIGADLKPIYPNPSYWNNSASVIFLDQPVGVSFGYCD----- 170
Qy 165 PYTVGVVNSSPAGVQGRPTIDATLIDTTNLAAEAWEILOGLSGLPSLDSRVQSKDFS 224
Db :
171 -----SKVSTDDAAKDVIIFLDLFFERFPHL-----RNNDHF 203
Qy 225 LMTESYGHHGYPAFFNFHFVEQNRIANGSVNGVQLNFNSLGIINGIIDALQAPYYEPA 284
Db :
204 ISGESYAGHYLPKIAH-----EIVVAHEDSFENLSSVLIGNGTDFPLTQYOYEPMA 256
Qy 285 VNNTYGIKAVNETVYNMKFANQMPCGD-----LISTCKQTNRTALADYALCA 334
Db :
257 CGEG-GYPAVLE-----PEDCLDMRNLPLCLSLVDRCYKSH-----SVFSCV 298
Qy 335 EATNMCRDNVEGSPYAFAGRGVYDIR--HPYDDP-----TPPSYNNKFLLAKDSVMDAIGV 387
Db :
299 LADRYCEQQITG-VYEKSGRNPYDIRSKAEADDGACYOEEIYISDYLNQEEVQRALGT 357
Qy 388 NINYTQS-NNDVYFAFOQTGDFVWPVFIEDLEEIALPVRVSLIYGDADYICNWFGQAV 446
Db :
358 DVSSFQGCSDVGIGFAFTGDGPSF-PHQYVALLDQDINVLIVAGDKOICNWLGNLAW 416
Qy 447 SLAANSQAQFRSAGYTPLKVGVE--YGETREYGNFSFRVYEAGHEVPPYQPIASIQ 504
Db :
417 TEKLEWRNYNEEKQVLRWTWKSEETDETIGTSYGLTYLRIYDAGHMVPHDOPENSILQ 476
Qy 505 LFNRIT 510
Db :
477 MVNSWI 482

RESULT 11
ABG93281
ID ABG93281 standard; Protein; 550 AA.
XX AC ABG93281;
XX DT 21-NOV-2002 (first entry)
XX

[illegible]

C. albicans BAX-associated protein fragment SEQ ID 520.

Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
neurodegeneration; cell death.

Candida albicans.
WO200264766-A2 .
22-AUG-2002..
21-DEC-2001; 2001WO-EP15398.
22-DEC-2000; 2000EP-0870318.
04-JAN-2001; 2001EP-0870002.
09-JAN-2001; 2001EP-0870003.
(JANC) JANSEN PHARM NV.
Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
WPI; 2002-667002/71.
N-PSDB; ABQ76547.

New isolated nucleic acid representing a synthetic Bax gene, useful as
medicament for treating, preventing and/or alleviating yeast or fungal
infections or proliferative disorders, and/or preventing apoptosis in
certain diseases -
Claim 36; Figure 2; 344pp; English.

This invention describes a novel nucleic acid representing a synthetic
Bax gene. The Bax gene of the invention is useful for identifying
Bax-resistant yeast or fungi, identifying, or obtaining and identifying
Candida spp. sequences that are differentially expressed in a pathway
eventually leading to programmed cell death or identifying inhibitors or
inhibitor sequences of Bax-induced cell death. The products of the
invention have cytostatic, fungicide, immunosuppressive, virucide and
vasotropic activity and can be used in vaccines or for gene therapy. The
isolated nucleic acids, polypeptides, pharmaceutical compositions,
antisense molecules and antibodies are useful as medicaments or in
preparing a medicament for treating, preventing and/or alleviating
diseases associated with yeast or fungi or proliferative disorders, such
as cancer, or for preventing apoptosis in certain organisms. The compounds
or poly peptides , or the genetically modified organism are useful for
preparing a medicament for modifying the endogenous flora of humans and
other mammals. The vaccine is useful for immunising against yeast or
fungal infections. Apoptosis-related diseases include autoimmune disease,
ischaemia, diseases related with viral infections or neurodegenerations.
This sequence represents a polypeptide associated with the Bax gene
described in the disclosure of the invention.

Sequence 550 AA;

Query Match 18.2%; Score 541.5; DB 23; Length 550;
Best_Local Similarity 31.3%; Pred. No. 1.5e-38;
Matches 150; Conservative 63; Mismatches 157; Indels 109; Gaps 17;

QY 67 VKSYSGYVD-TSPSEHTFFWFPEARHNDETPITLWLNGGPGSDSLIGFLFELGPCVNS 125
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 133 VKQVTGYLDISLDKHLFWFFESRNDPKNDPILWLNGGPGSCSSTGLFELGESSINK 192
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 126 IPDDYNPHSNWEVSNLLPLSOLPLGVGFYSFTVDGSINPTVGTVVENSFAGVQGRYPTI 185
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 193 TLHPVNPYPWSNNSAVIFLDQPVGVSFY-----TGCD----- 227
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 186 DATLIDTTNLAAEAAAEIIQLGFLSLGPSLDSRVQSDFSLWTESYGHYGPAPFNHFYEQ 245
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 228 ----VKNTLTAAKDVVVFLELFQKEP----QFLTNKTHIAGESVAGHYIPAF----- 272
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 186 DATLIDTTNLAAEAAAEIIQLGFLSLGPSLDSRVQSDFSLWTESYGHYGPAPFNHFYEQ 245
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 228 ----VKNTLTAAKDVVVFLELFQKEP----QFLTNKTHIAGESVAGHYIPAF----- 272
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 186 DATLIDTTNLAAEAAAEIIQLGFLSLGPSLDSRVQSDFSLWTESYGHYGPAPFNHFYEQ 245
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB 228 ----VKNTLTAAKDVVVFLELFQKEP----QFLTNKTHIAGESVAGHYIPAF----- 272
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||:

us-09-712-338-2.rag

Mon Nov 24 13:41:08 2003

PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143342.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159384.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			

Query Match

12.9%; Score 383; DB 21; Length 482;

PR	21-APR-1999;	99US-0130444
PR	23-APR-1999;	99US-0130510
PR	23-APR-1999;	99US-0130691
PR	28-APR-1999;	99US-0131449
PR	30-APR-1999;	99US-0132048
PR	30-APR-1999;	99US-0132407
PR	04-MAY-1999;	99US-0132484
PR	05-MAY-1999;	99US-0132485
PR	06-MAY-1999;	99US-0132486
PR	06-MAY-1999;	99US-0132487
PR	07-MAY-1999;	99US-0132863
PR	11-MAY-1999;	99US-0134256
PR	14-MAY-1999;	99US-0134218
PR	14-MAY-1999;	99US-0134219
PR	14-MAY-1999;	99US-0134221
PR	14-MAY-1999;	99US-0134370
PR	18-MAY-1999;	99US-0134768
PR	19-MAY-1999;	99US-0134941
PR	20-MAY-1999;	99US-0135124
PR	21-MAY-1999;	99US-0135353
PR	24-MAY-1999;	99US-0135629
PR	25-MAY-1999;	99US-0136021
PR	27-MAY-1999;	99US-0136392
PR	28-MAY-1999;	99US-0136782
PR	01-JUN-1999;	99US-0137222
PR	03-JUN-1999;	99US-0137528
PR	04-JUN-1999;	99US-0137502
PR	07-JUN-1999;	99US-0137724
PR	08-JUN-1999;	99US-0138094
PR	10-JUN-1999;	99US-0138540
PR	10-JUN-1999;	99US-0138847
PR	14-JUN-1999;	99US-0139119
PR	16-JUN-1999;	99US-0139432
PR	17-JUN-1999;	99US-0139453
PR	18-JUN-1999;	99US-0139459
PR	18-JUN-1999;	99US-0139459
PR	18-JUN-1999;	99US-0139461
PR	18-JUN-1999;	99US-0139462
PR	18-JUN-1999;	99US-0139463
PR	18-JUN-1999;	99US-0139750
PR	18-JUN-1999;	99US-0139763
PR	21-JUN-1999;	99US-0139817
PR	22-JUN-1999;	99US-0139899
PR	23-JUN-1999;	99US-0140353
PR	23-JUN-1999;	99US-0140354
PR	24-JUN-1999;	99US-0140695
PR	24-JUN-1999;	99US-0140623
PR	29-JUN-1999;	99US-0140991
PR	30-JUL-1999;	99US-0141287
PR	01-JUL-1999;	99US-0141842
PR	02-JUL-1999;	99US-0142154
PR	06-JUL-1999;	99US-0142055
PR	08-JUL-1999;	99US-0142390
PR	09-JUL-1999;	99US-0142803
PR	12-JUL-1999;	99US-0142920
PR	13-JUL-1999;	99US-0142977
PR	13-JUL-1999;	99US-0143542
PR	14-JUL-1999;	99US-0143624
PR	16-JUL-1999;	99US-0144005
PR	16-JUL-1999;	99US-0144085
PR	19-JUL-1999;	99US-0144325
PR	19-JUL-1999;	99US-0144331
PR	19-JUL-1999;	99US-0144332
PR	19-JUL-1999;	99US-0144333
PR	19-JUL-1999;	99US-0144335

PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145088.
PR	21-JUL-1999;	99US-0145088.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145122.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159283.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0144352.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 12.9%; Score 383; DB 21; Length 502;		
Best Local Similarity 26.2%; Pred. No. 1.le-24;		
Matches 135; Conservative 68; Mismatches 202; Indels 110; Gaps 18;		
QY	33	PKNPTGVKTLTTANNVTIRYKE-----PGAEGVCETTPG-----VKSYSGYVDTSPESHT 82
Db	52	PSFKGVSSSGDTSFHNVIQRESVSPKDKLIQLPGQSDVTFKQGGYVAVNKPAGR 111
QY	83	F-FWFFEARHNPETAPITLWNGFGSDSLIGLFEELGPGHVNSTFDP-YINPHSWNEV 139
Db	112	FLYYFVETIKPGNTTPLVWENGPGSCSLGAFKELGPPRVHSDGKTLFRNPYSWNE 171
QY	140	SNLLFLSQPLGVGFSYSDTVGDSINPVTGVWNSFAGVQGRYPTIDATLIDTTLAAEA 199
Db	172	ANVLFLETVPVGTGFSYS-----NSPINGQGDYAT-----AED 204
QY	200	AMEILQGLSLPLSDSRVQSKDFSLWTSYGHYGPAPFNHFYEQNERIANGSVNGVOL 259
Db	205	NYMFLVNWLERFPEYKGR---DIYTAGOSYAGHYVPQLAQIILHFN-----NOTLI 252
QY	260	NENSLGIINGIDEAIQAPYEPFVANNYTGKAVNETVYNMKFANQMPNGQDILSTC 319
Db	253	NLRGILIGNPSLNREIQDDFGYKMF--SHGLIS-OQOMDNVNF-----CTD--SDL 300
QY	320	KQTNRTALADYALCAEATNMCRDNVGPYAFAGRGVYDIRHPY-----363
Db	301	YDWDKCHLASQKTEAOKTHL-----DIYNIYAPLCNLSTLSSEPKCTTI 345
QY	364	--DDPTPPSYNFKFLAKDSVMDAIGN-----INYTSNNNDVYAFQQTGDFVWNPFI 416
Db	346	MKADPCSGNYLKAYLNKEVOEALHANTTKIPYEWTSCTNKLWENNEKDRYV--SLTPI 403
QY	417	LEBILALPVRSLLIYGDADYICNWFQGVASLANYSQAQPSAGYVTLKVNQVEYGET 476
Db	404	LOELMGKGVRLYNGVDVILVIFTSTLAVTKTMNLTVVKEWR-----PWFTGGHVGFFT 458
QY	477	REY-GNFSFTRVVEAGHEVYQYFIASLQLFNRTI 510
Db	459	EDYKGNLFTVTVKGAGHSVFTDQPIHALNIFTFSFI 493
RESULT 15		
ABR38843		
ID	ABR38843 standard; Protein; 455 AA.	
XX	ABR38843;	
XX	24-APR-2003 (first entry)	

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:01:36 ; Search time 39 Seconds
 (without alignments)
 3672.287 Million cell updates/sec

Title: US-09-712-338-2
 Perfect score: 555
 Sequence: 1 MRGYFSLVPLVAASWALP.....HTQSSVPLTATSMSSVGMA 555

Scoring table: OLIGO
 Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	454	81.8	555	3 Q96VZ9	Q96VZ9 aspergillus
2	10	1.8	460	3 Q12569	Q12569 absidia zyc
3	9	1.6	67	2 Q937G9	Q937G9 klebsiella
4	9	1.6	147	2 Q936V3	Q936V3 pseudomonas
5	9	1.6	405	2 Q66024	Q66024 pseudomonas
6	9	1.6	405	2 P71195	P71195 enterobacte
7	9	1.6	405	2 Q91UM7	Q91UM7 rhizobium m
8	9	1.6	405	2 Q56451	Q56451 xanthomonas
9	9	1.6	405	2 Q9F3W1	Q9F3W1 pseudomonas
10	9	1.6	468	10 Q9SHU3	Q9SHU3 arabidopsis
11	9	1.6	469	10 Q9FH06	Q9FH06 arabidopsis
12	9	1.6	473	10 Q9FH05	Q9FH05 arabidopsis
13	9	1.6	479	10 Q9MAR8	Q9MAR8 arabidopsis
14	9	1.6	495	2 Q48384	Q48384 klebsiella
15	9	1.6	2105	5 Q17679	Q17679 caenorhabdi
16	8	1.4	116	15 Q75691	Q75691 human immun

17	8	1.4	166	17	Q59236	O59236 pyrococcus
18	8	1.4	263	2	Q9ZF75	Q9ZF75 burkholderi
19	8	1.4	264	10	Q9LXY6	Q9LXY6 arabidopsis
20	8	1.4	264	10	Q8GVD2	Q8GVD2 zea mays tm
21	8	1.4	270	16	Q9RRC8	Q9RRC8 deinococcus
22	8	1.4	329	17	Q8TURS	Q8TURS methanopyru
23	8	1.4	362	10	Q9XEB1	Q9XEB1 sorghum bic
24	8	1.4	425	10	O82229	O82229 arabidopsis
25	8	1.4	425	10	O65568	O65568 arabidopsis
26	8	1.4	434	11	Q8BKV9	Q8BKV9 mus musculu
27	8	1.4	437	10	Q9FWG1	Q9FWG1 oryza sativ
28	8	1.4	437	11	Q9CZK1	Q9CZK1 mus musculu
29	8	1.4	437	11	Q9DBG5	Q9DBG5 mus musculu
30	8	1.4	447	10	Q8S8K6	Q8S8K6 arabidopsis
31	8	1.4	452	10	Q9ZQO0	Q9ZQO0 arabidopsis
32	8	1.4	456	10	Q9SV78	Q9SV78 arabidopsis
33	8	1.4	456	10	Q9C7E2	Q9C7E2 arabidopsis
34	8	1.4	458	10	O22803	O22803 arabidopsis
35	8	1.4	459	10	Q8RZS0	Q8RZS0 oryza sativ
36	8	1.4	459	10	Q9SFB5	Q9SFB5 arabidopsis
37	8	1.4	461	10	Q93Y09	Q93Y09 arabidopsis
38	8	1.4	465	10	O04084	O04084 arabidopsis
39	8	1.4	465	10	Q94K84	Q94K84 arabidopsis
40	8	1.4	465	10	Q8VY01	Q8VY01 arabidopsis
41	8	1.4	465	10	Q9M099	Q9M099 arabidopsis
42	8	1.4	470	10	O22732	O22732 arabidopsis
43	8	1.4	472	10	Q9LSM9	Q9LSM9 arabidopsis
44	8	1.4	474	11	Q9D2D1	Q9D2D1 mus musculu
45	8	1.4	479	10	Q949Q7	Q949Q7 arabidopsis

ALIGNMENTS

RESULT 1

ID	Q96VZ9	PRELIMINARY;	PRT;	555 AA.
AC	Q96VZ9			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DB	Carboxypeptidase S1 precursor (EC 3.4.16.6).			
GN	CPI.			
OS	Aspergillus oryzae.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=5062;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TK3;			
RA	van Den Broek P.;			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF394242; AAK77166.1; -			
DR	InterPro; IPR001563; Serine_carpept.			
DR	Pfam; PF00450; serine_carpept; 1.			
DR	PRINTS; PR00724; CRBOXYPTASEC.			
DR	ProDom; PD001189; Serine_carpept; 1.			
DR	PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.			
KW	Carboxypeptidase; Hydrolase; Signal.			
FT	SIGNAL 1 18			
SQ	SEQUENCE 555 AA; 61168 MW; 456B63B0CB55222 CRC64;			

Query Match 81.8%; Score 454; DB 3; Length 555;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MRGYFSLVPLVAASWALPGSTPASVGRQLPKNPTGVKTLTTANNVTIRKEPAGCV 60
Db	1	MRGYFSLVPLVAASWALPGSTPASVGRQLPKNPTGVKTLTTANNVTIRKEPAGCV 60
QY	61	CTTTGVKSYSGYVDTSPESHFTFFFFFAHNPETAPITLWNGFGSDSLGLFEELGP 120
Db	61	CTTTGVKSYSGYVDTSPESHFTFFFFFAHNPETAPITLWNGFGSDSLGLFEELGP 120

Db 430 RVEAGHEVP 439

RESULT 3
Q937G9 PRELIMINARY; PRT; 67 AA.

ID Q937G9
AC Q937G9
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative transposition protein (Fragment).
GN TNIO.
OS Klebsiella sp. LS13-39.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=143776;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS13-39; TRANSPOSON=Tn5056;
RC MEDLINE=21604134; PubMed=11763242;
RX Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA Yuriyeva O.V., Nikiforov V.G.;
RT "Mercury resistance transposons of Gram-negative environmental
bacteria and their classification.",
RL Res. Microbiol. 152:811-822(2001).
DR EMBL; AJ302770; CAC82957.1; --
FT NON_TER 1
SQ SEQUENCE 67 AA; 7547 MW; 3EECC0A1149C62D9 CRC64;

Query Match 1.6%; Score 9; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96
Db 3 EARNHPETA 11

RESULT 4
Q936V3 PRELIMINARY; PRT; 147 AA.

ID Q936V3
AC Q936V3
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative transposition protein TniQ (Fragment).
GN TNIO.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BW20; TRANSPOSON=Tn5053V4;
RC MEDLINE=21604134; PubMed=11763242;
RX Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA Yuriyeva O.V., Nikiforov V.G.;
RT "Mercury resistance transposons of Gram-negative environmental
bacteria and their classification.",
RL Res. Microbiol. 152:811-822(2001).
DR EMBL; AJ296313; CAC82974.1; --
FT NON_TER 1
SQ SEQUENCE 147 AA; 16500 MW; 24636FA7A1489ABD CRC64;

Query Match 1.6%; Score 9; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96
Db 83 EARNHPETA 91

QY 121 CHVNSTFDYINPWSNEVSNLLFLSQPLGVGFSYSYDVTGDSINPVTGWNSPAGVQG 180
Db 121 CHVNSTFDYINPWSNEVSNLLFLSQPLGVGFSYNDVDSINPVTGWNSPAGVQG 180
QY 181 RYPTIDATLIDTNLAEEAAMEILQGLSLDSRVQSKDFSLWTSYGGHYGPAPFN 240
Db 181 RYPTIDATLIDTNLAEEAAMEILQGLSLDSRVQSKDFSLWTSYGGHYGPAPFN 240
QY 241 HFEYQNERIANGSVNGVQLNFNSLGLINGIDEAICAPYYPFAVNNYTGKAVNETVYN 300
Db 241 HFEYQNERIANGSVNGVQLNFNSLGLINGIDEAICAPYYPFAVNNYTGKAVNETVYN 300
QY 301 YMKFANQMPNGQDILISTCKQNRNTALADYALCAEATNMCRDNVEGYPYAFAGRGVYDIR 360
Db 301 YMKFANQMPNGQDILISTCKQNRNTALADYALCAEATNMCRDNVEGYPYAFAGRGVYDIR 360
QY 361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNNYNTQSNNDVYVAFQOTGDFVWPNFIEDLEEI 420
Db 361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNNYNTQSNNDVYVAFQOTGDFVWPNFIEDLEEI 420
QY 421 LALPVRVSLIYGDADVICNWFQGVAVSLAANYSOAAQFRSAGYTPKLVNGVVEYGETREYG 480
Db 421 LALPVRVSLIYGDADVICNWFQGVAVSLAANYSOAAQFRSAGYTPKLVNGVVEYGETREYG 480
QY 481 NFSFTRVYAGHEVPYQPIASQLFNRTIFGWDIAEGQKKIWPSTKNGTATATHTQSS 540
Db 481 NFSFTRVYAGHEVPYQPIASQLFNRTIFGWDIAEGQKKIWPSTKNGTATATHTQSS 540
QY 541 VPLPTATSMSSVGMA 555
Db 541 VPLPTATSMSSVGMA 555

RESULT 2
Q12569 PRELIMINARY; PRT; 460 AA.

ID Q12569
AC Q12569
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Prepro-carboxypeptidase Z (EC 3.4.16.5).
GN SPCZ.
OS Absidia zychnae.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Absidia.
OX NCBI_TaxID=36079;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRIC 1199;
RC MEDLINE=95308538; PubMed=7788719;
RX Lee B.R., Takeuchi M., Kobayashi Y.;
RT "Molecular cloning and sequence analysis of the scpZ gene encoding the
serine carboxypeptidase of Absidia zychnae.",
RL Curr. Genet. 27:159-165(1995).
DR EMBL; D16519; BAA03966.1; --
DR HSSP; P00729; 1YSC.
DR MEROPS; S10.010; --
DR Interpro; IPR001563; Serine_carbpept.
DR Interpro; IPR000379; Serestr_site.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
DR Carboxypeptidase; Hydrolase.
FT CHAIN 52 460 MATURE CARBOXYPEPTIDASE Z.
SQ SEQUENCE 460 AA; 50825 MW; CE3A4107645488B6 CRC64;

Query Match 1.8%; Score 10; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 486 RVEAGHEVP 495
Db 486 RVEAGHEVP 495

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 EARNPETA 96
Db 341 EARNPETA 349

RESULT 7
Q91UM7
ID Q91UM7 PRELIMINARY; PRT; 405 AA.
AC Q91UM7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE ThiQ protein.
GN ThiQ.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSB102.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=mercury resistance transposon Tn5718;
RA Schneiker S., Keller M., Droegge M., Lanka E., Puhler A.,
RA Salbitschka W.;
RT "The genetic organization and evolution of the broad-host-range
RT mercury resistance plasmid pSB102 isolated from a microbial population
RT residing in the rhizosphere of alfalfa.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ304453; CAC79210.1;
KW Plasmid.
SQ SEQUENCE 405 AA; 45711 MW; 2396F2C89A2DE7B2 CRC64;

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 EARNPETA 96
Db 341 EARNPETA 349

RESULT 8
Q56451
ID Q56451 PRELIMINARY; PRT; 405 AA.
AC Q56451;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Mercury resistant transposon TNS053 (Central Asia).
GN ThiQ.
OS Xanthomonas sp.
OG Plasmid RPI.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=29446;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93253772; PubMed=8387603;
RA Kholodii G.Y.A., Yurieva O.V., Lomovskaya O.L., Gorlenko Zh.M.,
RA Mindlin S.Z., Nikiforov V.G.;
RT "TNS053, a mercury resistance transposon with integron's ends.";
RL J. Mol. Biol. 230:1103-1107(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96130850; PubMed=8594337;
RA Kholodii G.Y., Mindlin S.Z., Bass I.A., Yurieva O.V., Minakhina S.V.,
RA Nikiforov V.G.;
RT "Four genes, two ends, and a res region are involved in transposition
RT of Tn5053: a paradigm for a novel family of transposons carrying either
RT a mer operon or an integron.";
RL Mol. Microbiol. 17:1189-1200(1995).

```

```

RN RP SEQUENCE FROM N.A.
RA Kholodii G.Y.;
RL Russ. J. Genet. 31:1447-1451 (1995).
DR EMBL; L40595; AAA98331.1; -.
KW Plasmid.
SQ SEQUENCE 405 AA; 45830 MW; 89880020E32684C3 CRC64;

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96
DB 341 EARNHPETA 349

RESULT 9
ID Q9F3W1 PRELIMINARY; PRT; 405 AA.
AC Q9F3W1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TniQ protein.
GN TniQ.
OS Pseudomonas sp. ED23-33.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=83781;
RN [1]
RP SEQUENCE FROM N.A.
RA Minakina S., Minakhin L., Kholodii G., Mindlin S., Gorlenko Z.H.,
RA Yurieva O., Nikiforov V.;
RT "Molecular inventory of transposons from environmental bacteria:
RT epidemic dissemination of Tn21-, Tn5041-, and Tn5053-related
RT transposons."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y17897; CAC14709.1; -.
SQ SEQUENCE 405 AA; 45706 MW; 033A913AF44E43EE CRC64;

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96
DB 341 EARNHPETA 349

RESULT 10
ID Q9SHU3 PRELIMINARY; PRT; 468 AA.
AC Q9SHU3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative serine carboxypeptidase II.
GN Ar2G12480.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.I., Barnstead M.B., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,

RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Lin X.;
RC STRAIN=cv. Columbia;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007268; AAD28662.1; -.
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR002000; Lamp.
DR InterPro; IPR001563; Serine carboxypept.
DR Pfam; PF00450; serine carboxypept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxypept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00310; LAMP_1; 1.
KW Carboxypeptidase.
SQ SEQUENCE 468 AA; 52836 MW; 47DD2055AB9AE1D7 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 468;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPG 107
DB 77 TLWLNGGPG 85

RESULT 11
ID Q9FH06 PRELIMINARY; PRT; 469 AA.
AC Q9FH06;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serine carboxypeptidase-II like.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63 (2000).
DR EMBL; AB023032; BAB10196.1; -.
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR001563; Serine carboxypept.
DR Pfam; PF00450; serine carboxypept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxypept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 469 AA; 52541 MW; 90357E27B9B64527 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPG 107
DB 74 TLWLNGGPG 82

```

```

RESULT 12
Q9FH05 PRELIMINARY; PRT; 473 AA.
AC Q9FH05;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Serine carboxypeptidase II-like (AT5942240/K5J14_4).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
RA Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
RA Lin J., Liu S.X., Miranda M., Narusaka M., Narusaka M.,
RA Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
RA Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
RA Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023032; BAB10197.1; -
DR EMBL; AF361604; AAK32772.1; -
DR EMBL; AY143899; AAN28838.1; -
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -
DR InterPro; IPR001563; Serine carboxypeptidase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine_carboxypeptidase.
DR PRINTS; PR00724; CRBOXPTASEC.
DR ProDom; PD001189; Serine carboxypeptidase.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 473 AA; 52874 MW; 714E0F29BC251590 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 473;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 TLWINGGPG 107
Db 78 TLWINGGPG 86

RESULT 13
Q9MAR8 PRELIMINARY; PRT; 479 AA.
AC Q9MAR8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Putative serine carboxypeptidases.
GN F28H19.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chlou J., Choi E., Gonzalez A.,
RA Howling B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaynsberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006423; AAF63101.1; -
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -
DR InterPro; IPR001563; Serine carboxypeptidase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine_carboxypeptidase.
DR PRINTS; PR00724; CRBOXPTASEC.
DR ProDom; PD001189; Serine carboxypeptidase.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 479 AA; 53922 MW; 87A36BB1C46E1080 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 TLWINGGPG 107
Db 83 TLWINGGPG 91

RESULT 14
Q48384 PRELIMINARY; PRT; 495 AA.
AC Q48384;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE ORF6 protein.
OS Klebsiella aerogenes.
OG Klebsiella aerogenes.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=28451;
RN [1]
RP SEQUENCE FROM N.A.
RA Sundstroem L.;
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94352994; PubMed=8195081;
RA Radstrom P., Skold O., Swedberg G., Flensburg J., Roy P.H.,
RA Sundstrom L.;
RT "Transposon Tn5090 of plasmid R751, which carries an integron, is
RT related to Tn7, Mu, and the retroelements."
RL J. Bacteriol. 176:3257-3268 (1994).
KW Plasmid.
SQ SEQUENCE 495 AA; 55107 MW; A36EF0E3F3B70B13 CRC64;

Query Match 1.6%; Score 9; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.8;

```

Mon Nov 24 13:41:08 2003

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96
 |||||
 DB 431 EARNHPETA 439

RESULT 15

Q17679 PRELIMINARY; PRT; 2105 AA.
 AC Q17679;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Y16B4A.2 protein.
 GN Y16B4A.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70203; CAA94110.1; -.
 DR EMBL; AL023825; CAA94110.1; JOINED.
 DR EMBL; Z93339; CAA94110.1; JOINED.
 DR EMBL; Z93339; CAB07544.1; -.
 DR EMBL; AL023825; CAB07544.1; JOINED.
 DR EMBL; Z70203; CAB07544.1; JOINED.
 DR EMBL; AL023825; CAA19443.1; -.
 DR EMBL; Z93339; CAA19443.1; JOINED.
 DR EMBL; Z70203; CAA19443.1; JOINED.
 DR HSP; P10619; LIVY.
 DR WormPep; Y16B4A.2; CE21374.
 DR InterPro; IPR001563; Serine_carbpept.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR Pfam; PF00450; serine_carbpept; 4.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine_carbpept; 6.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 4.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 2.
 SQ SEQUENCE 2105 AA; 234335 MW; 139CB6E96D75B3FB CRC64;

Query Match 1.6%; Score 9; DB 5; Length 2105;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPGS 108
 |||||
 DB 549 LWLNGGPGS 557

Search completed: November 21, 2003, 18:05:16
 Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:01:06 ; Search time 17 Seconds
(without alignments)
1535.284 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 555

Sequence: 1 MRGYEFLSVLPVLAASWALP.....HTQSSVPLPTATSMSSVGMNA 555

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	1.8	423	1	CPS1_PENJA
2	9	1.6	508	1	YBY9_YEAST
3	8	1.4	36	1	PSBY_ODOSI
4	8	1.4	423	1	CBP2_WHEAT
5	8	1.4	436	1	CP22_HORVU
6	8	1.4	470	1	YSS2_CAEEL
7	8	1.4	474	1	PRTP_MOUSE
8	8	1.4	476	1	CBP2_HORVU
9	8	1.4	480	1	PRTP_HUMAN
10	8	1.4	499	1	CBP1_HORVU
11	8	1.4	510	1	CBP1_ORYSA
12	8	1.4	516	1	CP23_HORVU
13	8	1.4	523	1	CBPY_PICPA
14	8	1.4	532	1	CBPY_YEAST
15	8	1.4	542	1	CBPY_CANAL
16	8	1.4	574	1	YPP3_CAEEL
17	8	1.4	1002	1	CBPY_SCHPO
18	8	1.4	4349	1	FAT2_HUMAN
19	7	1.3	25	1	VCX8_ODOSI
20	7	1.3	110	1	RLA4_YEAST
21	7	1.3	117	1	ARR2_ECOLI
22	7	1.3	158	1	GML_HUMAN
23	7	1.3	216	1	OL1F_HUMAN
24	7	1.3	216	1	OL1I_HUMAN
25	7	1.3	225	1	SOC1_MOUSE
26	7	1.3	250	1	PSD8_CAEEL
27	7	1.3	290	1	SUHE MYCTO
28	7	1.3	297	1	VGLG_HRSV3
29	7	1.3	298	1	VGLG_HRSV3
30	7	1.3	298	1	VGLG_HRSV4
31	7	1.3	315	1	O3A1_HUMAN
32	7	1.3	315	1	O3A3_HUMAN
33	7	1.3	317	1	PLC_LISMO

34 7 1.3 321 1 STILL_DROME
35 7 1.3 334 1 ADD_VIBCH
36 7 1.3 341 1 GLK_RHILO
37 7 1.3 344 1 DPO4_STRPN
38 7 1.3 346 1 PA6A_MOUSE
39 7 1.3 353 1 Y740_AQUAE
40 7 1.3 354 1 DP42_RHIME
41 7 1.3 356 1 DPO4_STAAM
42 7 1.3 356 1 DPO4_STAAM
43 7 1.3 356 1 E13L_TOBAC
44 7 1.3 356 1 MURG_YERPE
45 7 1.3 357 1 DP42_AGRTS

ALIGNMENTS

RESULT 1
CPS1_PENJA
AC P34946;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase S1 (EC 3.4.16.6)
OS Penicillium janthinellum (Penicillium vitale).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5079;
RN [1]
RP SEQUENCE.
RX MEDLINE=94039747; PubMed=8224168;
RA Svendsen I., Hofmann T., Endrizzi J., Remington S.J., Breddam K.;
RT "The primary structure of carboxypeptidase S1 from Penicillium janthinellum";
RL FEBS Lett. 333:39-43(1993).
CC -|- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine or lysine residue.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
DR PIR; S38953; S38953.
DR HSSP; P00729; LYSC.
DR MEROPS; S10.008; -.
DR InterPro; IPR000379; Ser_estrs site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein.
FT ACT_SITE 143 143 POTENTIAL.
FT ACT_SITE 340 340 BY SIMILARITY.
FT ACT_SITE 397 397 BY SIMILARITY.
FT BINDING 343 343 SUBSTRATE.
FT BINDING 398 398 SUBSTRATE.
FT DISULFID 8 68
FT DISULFID 55 300
FT DISULFID 223 246
FT DISULFID 230 239
FT CARBOHYD 200 200 N-LINKED (GLCNAC...),
SQ SEQUENCE 423 AA; 46496 MW; 55F3333608B12729 CRC64;

Query Match 1.8%; Score 10; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.046; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 227 TSYGCHYCP 236

Db 141 TSYGCHYCP 150

RESULT 2

YBY9_YEAST

ID YB9 YEAST STANDARD; PRT; 508 AA.
 AC P38109;
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative serine carboxypeptidase in ESRI-IRAI intergenic region
 DE (EC 3.4.16.-)
 GN YBR139W OR YBR1015.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=94378717; PubMed=8091856;
 RA Becam A.-M., Cullin C., Grzybowska E., Lacroute F., Nasr F.,
 RA Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
 RA Herbert C.J.;
 RT "The sequence of 29.7 kb from the right arm of chromosome II reveals
 RT 13 complete open reading frames, of which ten correspond to new
 RT genes.";
 RL Yeast 10:S1-S11(1994).
 RN [2]
 RP DISCUSSION OF SEQUENCE.
 RC MEDLINE=95042830; PubMed=7954890;
 RA Nasr F., Becam A.-M., Grzybowska E., Zagulski M., Slonimski P.P.,
 RA Herbert C.J.;
 RT "An analysis of the sequence of part of the right arm of chromosome
 RT II of S. cerevisiae reveals new genes encoding an amino-acid permease
 RT and a carboxypeptidase.";
 RL Curr. Genet. 26:1-7(1994).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X75891; CAA53497.1; -;
 DR EMBL; Z36008; CAA85097.1; -;
 DR PIR; S46008; S46008.
 DR HSSP; P00729; 1CPY.
 DR MEROPS; S10 UPW; -;
 DR SGD; S0000343; YBR139W
 DR InterPro; IPR000379; Ser_estr_ site.
 DR InterPro; IPR001563; Serine_carbpept.
 DR Pfam; PF00450; serine_carbpept; 1.
 DR PRINTS; PR00724; CRBOXVPTASEC.
 DR ProDom; PD001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 KW Hypothetical protein; Hydrolase; Carboxypeptidase.
 FT ACT_SITE 219 BY SIMILARITY.
 FT ACT_SITE 219 BY SIMILARITY.
 FT ACT_SITE 415 BY SIMILARITY.
 FT ACT_SITE 474 BY SIMILARITY.
 FT ACT_SITE 474 BY SIMILARITY.
 SQ SEQUENCE 508 AA; 57639 MW; AAB2806C8EE2ED81 CRC64;

 Query Match 1.6%; Score 9; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 147 QPLGVGFPSY 155
 Db 171 QPLGVGFPSY 179

 RESULT 3
 PSBY ODOSI
 ID PSBY_ODOSI STANDARD; PRT; 36 AA.

P49543;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Photosystem II protein Y.
 DE (PSBY-A OR PSBY-1 OR YCF32.1) AND (PSBY-B OR PSBY-2 OR YCF32.2).
 GN Odontella sinensis (Marine centric diatom).
 OS Chloroplast.
 OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
 OC Bidulphiophycidae; Eupodiscales; Eupodisaceae; Odontella.
 OX NCBI_TaxID=2839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
 RT "The chloroplast genome of a chlorophyll a+c-containing alga,
 RT Odontella sinensis.";
 RL Plant Mol. Biol. Rep. 13:336-342(1995).
 CC -!- FUNCTION: MANGANESE-BINDING POLYPEPTIDE WITH L-ARGININE
 CC METABOLIZING ENZYME ACTIVITY. COMPONENT OF THE CORE OF PHOTOSYSTEM
 CC II (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
 CC -!- SIMILARITY: BELONGS TO THE PSBY FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z67753; CAA91612.1; -;
 DR EMBL; Z67753; CAA91666.1; -;
 DR PIR; S78239; S78239.
 KW Photosystem II; Chloroplast; Transmembrane; Thylakoid.
 FT DOMAIN 1 6 LUMENAL (POTENTIAL).
 FT TRANSMEM 7 23 POTENTIAL.
 FT DOMAIN 24 36 STROMAL (POTENTIAL).
 FT SEQUENCE 36 AA; 3921 MW; 22989D3E7DAE2B63 CRC64;

 Query Match 1.4%; Score 8; DB 1; Length 36;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 12 LVAASWAL 19
 Db 13 LVAASWAL 20

 RESULT 4
 ID CBP2 WHEAT STANDARD; PRT; 423 AA.
 AC P08819;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine carboxypeptidase II chains A and B (EC 3.4.16.6)
 DE (Carboxypeptidase D) (CPDW-II) (CP-WII).
 GN CBP2.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE.
 RA Breddam K., Soerensen S.B., Svendsen I.;
 RT "Primary structure and enzymatic properties of carboxypeptidase II
 RT from wheat bran.";
 RL Carlsberg Res. Commun. 52:297-311(1987).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).
 RX MEDLINE=90216664; PubMed=2324089;

RA Liao D.-I., Remington S.J.;
 RT "Structure of wheat serine carboxypeptidase II at 3.5-A resolution. A
 RL new class of serine proteinase.";
 RN J. Biol. Chem. 265:6528-6531(1990).
 RP [3]
 RX X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RA MEDLINE=93003139; PubMed=1390755;
 RA Liao D.-I., Breddam K., Sweet R.M., Bullock T., Remington S.J.;
 RT "Refined atomic model of wheat serine carboxypeptidase II at 2.2-A
 RT resolution.";
 RL Biochemistry 31:9796-9812(1992).
 CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
 CC or lysine residue.
 CC -!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
 CC COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 DR PIR; A29639; A29639.
 DR PDB; 3SC2; 31-OCT-93.
 DR PDB; 1WHS; 22-JUN-94.
 DR PDB; 1WHT; 22-JUN-94.
 DR PDB; 1BCR; 08-MAR-96.
 DR PDB; 1BCS; 08-MAR-96.
 DR MEROPS; S10.005;
 DR InterPro; IPR000379; Ser esters_site.
 DR InterPro; IPR001563; Serine_carbpept.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; 3D-structure.
 FT CHAIN 1 263
 FT NON_CONS 263 264
 FT CHAIN 264 423
 FT ACT_SITE 158 158
 FT ACT_SITE 340 340
 FT ACT_SITE 392 392
 FT DISULFID 65 303
 FT DISULFID 222 234
 FT DISULFID 258 270
 FT CARBOHYD 116 116
 FT CARBOHYD 127 127
 FT CARBOHYD 259 259
 FT CARBOHYD 291 291
 FT CARBOHYD 297 297
 FT CARBOHYD 421 421
 FT VARIANT 1 3
 FT HELIX 6 9
 FT STRAND 10 10
 FT TURN 15 16
 FT TURN 24 32
 FT TURN 33 36
 FT STRAND 37 44
 FT HELIX 48 50
 FT STRAND 55 59
 FT TURN 62 64
 FT STRAND 65 65
 FT TURN 67 70
 FT HELIX 71 74
 FT STRAND 79 81
 FT HELIX 83 85
 FT STRAND 88 90
 FT TURN 92 93
 FT HELIX 95 97
 FT TURN 98 98
 FT STRAND 100 104
 FT TURN 108 109
 FT TURN 111 112
 FT STRAND 114 115
 FT HELIX 118 122
 FT HELIX 126 143
 FT HELIX 145 147
 FT TURN 148 149
 FT STRAND 151 157

N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 N-LINKED (GLCNAC. . .)
 MISSING (IN A' CHAIN).

INTERCHAIN.
 SERINE CARBOXYPEPTIDASE II, CHAIN B.
 SERINE CARBOXYPEPTIDASE II, CHAIN A.

FT TURN 158 159
 FT HELIX 160 174
 FT TURN 175 175
 FT TURN 177 178
 FT STRAND 180 187
 FT STRAND 192 192
 FT HELIX 193 205
 FT TURN 206 208
 FT HELIX 212 222
 FT TURN 223 224
 FT HELIX 232 245
 FT TURN 250 251
 FT TURN 253 254
 FT TURN 269 270
 FT HELIX 271 279
 FT TURN 280 280
 FT HELIX 282 288
 FT TURN 289 289
 FT TURN 292 293
 FT STRAND 303 303
 FT HELIX 305 309
 FT TURN 310 310
 FT STRAND 317 317
 FT HELIX 319 327
 FT TURN 328 329
 FT STRAND 331 337
 FT TURN 338 339
 FT HELIX 345 353
 FT TURN 354 356
 FT STRAND 359 368
 FT TURN 369 370
 FT STRAND 371 379
 FT TURN 380 381
 FT STRAND 382 387
 FT TURN 388 389
 FT HELIX 394 397
 FT HELIX 399 411
 FT TURN 412 412
 SQ SEQUENCE 423 AA; 47111 MW; C64815BFF62B085C CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 423;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LWLNGGPG 107
 DB 57 LWLNGGPG 64
 RESULT 5
 CP22_HORVU
 ID_CP22_HORVU STANDARD; PRT; 436 AA.
 AC P55748;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine carboxypeptidase II-2 precursor (EC 3.4.16.6) (CP-MII.2)
 DE (Fragment).
 GN CXP;2-2.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxID=4513;
 OX [1]
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP STRAIN=cv. Alexis; TISSUE=Grain;
 RX MEDLINE=94336715; PubMed=7520177;
 RA Dal Degan F., Rocher A., Cameron-Mills V., von Wettstein D.;
 RT "The expression of serine carboxypeptidases during maturation and
 RT germination of the barley grain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213(1994).
 CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine

or lysine residue.

-! SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).

-! DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. LOW LEVELS IN THE DEVELOPING ALEURONE AND EMBRYO. ALSO FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.

-! PTM: THE LINKER PEPTIDE IS ENDOPEPTIDOLYTICALLY EXCISED DURING ENZYME MATURATION (BY SIMILARITY).

-! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to licenses@sib.ch).

CC
DR EMBL; X78878; CAB59202.1; -.

DR HSP: P08819; LWHT.
DR InterPro: IPR00379; Ser estrs site.
DR InterPro: IPR001563; Serine carboxyl.
DR Pfam: PF00450; serine Carboxyl; 1.
DR NCBI: P08819; carboxylase, ser.

DR EKINLIS; P000724; CRKOA1F1456C.
DR ProDom; PD001189; Serine carboxpt; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Multigene family.

FT	NON TER	1	1
FT CHAIN	<1	256	SERINE CARBOXYPEPTIDASE II-2, CHAIN A.
FT PROPP	257	270	LINKER PEPTIDE (BY SIMILARITY).
FT CHAIN	271	436	SERINE CARBOXYPEPTIDASE II-2, CHAIN B.
FT ACCT SIMP	149	149	BY SIMILARITY.

FT	217	BY SIMILARITY.
FT	229	BY SIMILARITY.
DISULFID	217	BY SIMILARITY.
ACT SITE	56	INTERCHAIN (BY SIMILARITY).
FT	403	BY SIMILARITY.
FT	350	BY SIMILARITY.
ACT SITE	350	BY SIMILARITY.
FT	350	BY SIMILARITY.

Query Match 1.4%; Score 8; DB 1; Length 436;
SQ SEQUENCE 436 AA; 48952 MW; E0F82D97E0C34DC9 CRC64;
FT CARBOHYD 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 253 BY SIMILARITY.
FT 281

```
Qy      100 LWLNGGPG 107
          L
          I
          V
          V
          V
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 48 LWLNGPG 55

RESULT 6

```

YSS2 CAEEL
ID_YSS2 CAEEL          PRT;  470 AA.
AC Q09991;
DT 01-FEB-1996 (Rel. 33, Created)
          (Rel. 33, Updated)

```

DT 01-FEB-1996 (rel. 33, last sequence update)
 DT 28-FEB-2003 (rel. 41, last annotation update)
 DE Putative serine carboxypeptidase K10B2.2 precursor (EC 3.4.16.-).
 GN K10B2.2.
 OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N A
DD

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RA Miller N.;
 RC STRAIN=Bristol N2;
 RE sequence no. 1.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

EMBL; U28730; AAC68259.1; -.
DR DR
PIR; T16606; T16606.
DR DR
HSSP; P10619; LIY.

DR MEROPS; S10.002; -.
DR WormPep; K10B2.2; CE02009.
DR InterPro; IPR000379; Ser_estrs_site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR

DR PRINTS; PRO0724; CRBOXYPTASEC.
DR DR
DR ProDom; PD001189; serine carboxpt; 2.
DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.

vibrio cholerae toxin udu-
vibriocidal protein udu-
glycomrotein:
carboxypeptidase:

KW	1	19	POTENTIAL.
Hypotnetical protein; hydrolase; carboxypeptidase; carboxypeptidase,			
Signal.			
FT SIGNAL	20	470	PUTATIVE SERINE CARBOXYPEPTIDASE K10B2.2.
FT CHAIN			
FT ACT SITE	169	169	BY SIMILARITY.

BY SIMILARITY.		BY SIMILARITY.	
FT	ACT SITE	380	380
FT	ACT SITE	441	441
FT	CARBOHYD	132	132
FT	CARBOHYD	316	316
FT	CARBOHYD	206	206
FT	CARBOHYD	206	206

SQ **CAREGRID** **FI**

SEQUENCE	470 AA;	53158 MW;	CCC2DACB75EF30FC CRC64;
Query Match	1.4%;	Score 8;	DB 1; Length 470;
Best Local Similarity	100.0%;	Pred No.	5.9;

	Matches	8;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	100	LWLNGPG	107							
Dh	71	LWLNGPG	78							

RESULT 7
PRTP MOUSE

ID	PRTP_MOUSE	STANDARD;	PAT;	4/4 AA.
AC	P16675; Q8VEF6;			
DT	01-AUG-1990	(Rel. 15, Created)		
DT	01-AUG-1990	(Rel. 15, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		

DE Lysosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A)
DE (Carboxypeptidase C) (Protective protein for beta-galactosidase).
DE PPGE.
OS *Mus musculus* (Mouse).
OS Chordata, Cephalata, Vertebrata, Euteleostomi.

CC Eukaryota; Metazoa;
OC Mammalia; Eutheria;
OX NCHI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=90170982; PubMed=2105523;
RA Galatz N.J., Gilleman N., Meijer D., D'Azzo A.;
RT "Mouse protective protein". cDNA cloning, sequence comparison, and
RT expression.";
DI J Biol Chem 265:4578-4584(1990)

U. Biol. Chem. 263:4076-4084(1988).
[2]
SEQUENCE FROM N.A.
RN
RN
RP
RX
RX
RA
RA

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Staelens M., Soares M.B., Ronald M.F., Casavant T.L., Scheetz T.E.,
RA

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA

RA Villaion D.K., Muzny D.M., Sodergren E.O., Lu A., Gibbs K.A.,

FAHEY J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 Buterfield V.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: PROTECTIVE PROTEIN APPEARS TO BE ESSENTIAL FOR BOTH THE
 CC ACTIVITY OF BETA-GALACTOSIDASE AND NEURAMINIDASE, IT ASSOCIATES
 CC WITH THESE ENZYMES AND EXERTS A PROTECTIVE FUNCTION NECESSARY FOR
 CC THEIR STABILITY AND ACTIVITY. THIS PROTEIN IS ALSO A
 CC CARBOXYPEPTIDASE AND CAN DEAMIDATE TACHYKININS.
 CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 CC broad specificity.
 CC -!- SUBUNIT: HETERODIMER OF A 32 kDa AND 20 kDa CHAINS HELD TOGETHER
 CC BY DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J05261; AAA39982.1; -;
 CC EMBL; BC018534; AAH18534.1; -;
 CC PIR; A35732; A35732.
 CC HSPSP; P10619; IIVY.
 CC MEROPS; S10.002; -;
 CC MGD; MG1.97748; Pgcb.
 CC InterPro; IPR00379; Ser. esters site.
 CC Pfam; PF001563; Serine_carbpept.
 CC PRINTS; PR00724; CRBOXYPTASEC.
 CC ProDom; PD001189; Serine_carbpept; 2.
 CC PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 CC PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 CC Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal; Lysosome.
 CC SIGNAL 1 23
 CC CHAIN 24 474 LYSOSOMAL PROTECTIVE PROTEIN.
 CC CHAIN 24 320 32 kDa CHAIN.
 CC CHAIN 321 474 20 kDa CHAIN.
 CC ACT_SITE 173 173 BY SIMILARITY.
 CC ACT_SITE 394 394 BY SIMILARITY.
 CC ACT_SITE 451 451 BY SIMILARITY.
 CC ACT_SITE 451 451 BY SIMILARITY.
 CC ACT_SITE 451 451 BY SIMILARITY.
 CC ACT_SITE 235 251 BY SIMILARITY.
 CC ACT_SITE 236 241 BY SIMILARITY.
 CC ACT_SITE 276 325 BY SIMILARITY.
 CC CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 425 425 G -> W (IN REF. 2).
 CC SEQUENCE 474 AA; 53844 MW; 8F4D5A7F4F5E6CE CRC64;
 SQ
 Query Match 1.4%; Score 8; DB 1; Length 474;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LWLNGGPG 107
 |||||
 75 LWLNGGPG 82
 Db
 RESULT 8
 CBP2_HORVU STANDARD; PRT; 476 AA.
 AC P08818; P93177;
 DT 01-NOV-1988 (Rel. 09, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine carboxypeptidase II precursor (EC 3.4.16.6) (Carboxypeptidase
 DE D) (CP-MII) [Contains: Serine carboxypeptidase II chain A; Serine
 DE carboxypeptidase II chain B].
 GN CBP2 OR CXP;2
 OS Hordeum vulgare (Barley).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Igri; TISSUE=Etiolated leaf;
 RA Rocher A., Lok F., Cameron-Mills V., von Wettstein D.;
 RT "The gene family of serine carboxypeptidases in barley.";
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC [2]
 CC SEQUENCE OF 35-294 AND 314-472.
 CC Soerensen S.B., Svendsen I., Breddam K.;
 CC "Primary structure of carboxypeptidase II from malted barley.";
 CC Carlsberg Res. Commun. 52:285-295(1987).
 CC -!- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF SMALL PEPTIDES (2-
 CC 5 RESIDUES) OR IN THE DEGRADATION OF STORAGE PROTEINS IN THE
 CC EMBRYO.
 CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
 CC or lysine residue.
 CC -!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
 CC COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND.
 CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.
 CC -!- DEVELOPMENTAL STAGE: SIMULTANEOUSLY PRESENT IN ALEURONE AND
 CC ENDOSPERM BETWEEN 20 AND 30 DAYS POSTANTHESIS. ACCUMULATES IN THE
 CC DEVELOPING GRAIN AND IS STORED IN ITS ACTIVE FORM IN THE MATURE
 CC GRAIN. ALSO FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y09602; CAA70815.1; -;
 CC PIR; T05701; T05701.
 CC HSPSP; P08819; LWHT.
 CC MEROPS; S10.005; -;
 CC InterPro; IPR001563; Serine_carbpept.
 CC Pfam; PF00450; serine_carbpept; 1.
 CC PRINTS; PR00724; CRBOXYPTASEC.
 CC ProDom; PD001189; Serine_carbpept; 1.
 CC PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 CC PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 CC Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal;
 CC Multigene family.
 CC SIGNAL 1 34
 CC CHAIN 35 294 SERINE CARBOXYPEPTIDASE II, CHAIN A.
 CC CHAIN 35 294 LINKER PEPTIDE (BY SIMILARITY).
 CC CHAIN 313 313 SERINE CARBOXYPEPTIDASE II, CHAIN B.
 CC ACT_SITE 314 476 BY SIMILARITY.
 CC ACT_SITE 190 190 BY SIMILARITY.
 CC ACT_SITE 390 390 BY SIMILARITY.
 CC ACT_SITE 443 443 BY SIMILARITY.
 CC ACT_SITE 97 353 INTERCHAIN (BY SIMILARITY).
 CC ACT_SITE 254 266 BY SIMILARITY.
 CC ACT_SITE 290 320 BY SIMILARITY.
 CC ACT_SITE 148 148 BY SIMILARITY.
 CC CARBOHYD 159 159 N-LINKED (GLCNAC. . .).
 CC CARBOHYD 291 291 N-LINKED (GLCNAC. . .).
 CC CARBOHYD 341 341 N-LINKED (GLCNAC. . .).
 CC CARBOHYD 347 347 N-LINKED (GLCNAC. . .).
 CC CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (PARTIAL).
 CC CARBOHYD 472 472 N-LINKED (GLCNAC. . .).

FT MOD_RES 314 314 BLOCKED.
 FT VARIANT 351 352 TN -> AT.
 FT CONFLICT 181 181 Y -> R (IN REF. 2).
 SQ SEQUENCE 476 AA; 52625 MW; 18685725B1A6B5E4 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 476;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LWINGPG 107
 Db 89 LWINGPG 96
 RESULT 9
 PRTP_HUMAN STANDARD; PRT; 480 AA.
 AC P10619; Q96KJ2; Q9BW69;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lysosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A)
 DE (Carboxypeptidase C) (Protective protein for beta-galactosidase).
 GN PPG8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=88311078; PubMed=3136930;
 RA Galjart N.J., Gillemans N., Harris A., van de Horst G.T.J.,
 RA Verheijen F.W., Galjaard H., D'Azzo A.,
 RT "Expression of cDNA encoding the human 'protective protein'
 RT associated with lysosomal beta-galactosidase and neuraminidase:
 RT homology to yeast proteases.";
 RL Cell 54:755-764 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharshaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McClay K., McMurray A.A.,
 RA Malne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Thorpe A.,
 RA Whitehead S.I., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 29-53 AND 327-351.
 RC TISSUE=Platelet;
 RX MEDLINE=90293074; PubMed=1694176;
 RA Jackman H.L., Tan F., Tamei H., Beuriling-Harbury C., Li X.-Y.,
 RA Skidgel R.A., Erdoes E.G.;
 RT "A peptidase in human platelets that deamidates tachykinins. Probable
 RT identity with the lysosomal 'protective protein'.";
 RL J. Biol. Chem. 265:11265-11272 (1990).
 RN [5]
 RP FUNCTION, AND MUTAGENESIS.
 RX MEDLINE=91117848; PubMed=1907282;
 RA Galjart N.J., Morreau H., Willemsen R., Gillemans N., Bonten E.J.,
 RA D'Azzo A.;
 RT "Human lysosomal protective protein has cathepsin A-like activity
 RT distinct from its protective function.";
 RL J. Biol. Chem. 266:14754-14762 (1991).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=96164441; PubMed=8591035;
 RA Rudenko G., Bonten E., D'Azzo A., Hol W.G.J.;
 RT "Three-dimensional structure of the human 'protective protein':
 RT structure of the precursor form suggests a complex activation
 RT mechanism.";
 RL Structure 3:1249-1259 (1995).
 RN [7]
 RP VARIANT GALACTOSIALIDOSIS VAL-440.
 RX MEDLINE=92097522; PubMed=1756715;
 RA Zhou X.Y., Galjart N.J., Willemsen R., Gillemans N., Galjaard H.,
 RA D'Azzo A.;
 RT "A mutation in a mild form of galactosialidosis impairs dimerization
 RT of the protective protein and renders it unstable.";
 RL EMBO J. 10:4041-4048 (1991).
 RN [8]
 RP VARIANTS GALACTOSIALIDOSIS ARG-49; ARG-65; LEU-90; ASN-249 AND
 RP CYS-395.
 RX MEDLINE=93293970; PubMed=8514852;
 RA Shimoto M., Fukuhara Y., Itoh K., Oshima A., Sakuraba H., Suzuki Y.;
 RT "Protective protein gene mutations in galactosialidosis.";
 RL J. Clin. Invest. 91:2393-2398 (1993).
 CC -! FUNCTION: PROTECTIVE PROTEIN APPEARS TO BE ESSENTIAL FOR BOTH THE
 CC ACTIVITY OF BETA-GALACTOSIDASE AND NEURAMINIDASE, IT ASSOCIATES
 CC WITH THESE ENZYMES AND EXERTS A PROTECTIVE FUNCTION NECESSARY FOR
 CC THEIR STABILITY AND ACTIVITY. THIS PROTEIN IS ALSO A
 CC CARBOXYPEPTIDASE AND CAN DEAMIDATE TACHYKININS.
 CC -! CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 CC broad specificity.
 CC -! SUBUNIT: HETERODIMER OF A 32 kDa AND 20 kDa CHAINS HELD TOGETHER
 CC BY DISULFIDE BONDS.
 CC -! SUBCELLULAR LOCATION: Lysosomal.
 CC -! DISEASE: Defects in PPG8 are the cause of galactosialidosis
 CC [MIM:256540], an autosomal recessive disease.
 CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M22960; AAA36476.1; --
 EMBL; AL008726; AAA15501.1; --
 EMBL; BC000597; AAH00597.1; --
 PIR; A31589; A31589.
 PDB; LIVY; 2I-APR-97.
 MEROPS; S10.002; --
 Genew; HGNC:9251; PPG8.
 MIM; 256540; --
 GO; GO:0005783; C:cytoplasmic reticulum; TAS.
 GO; GO:0004180; F:carboxypeptidase activity; TAS.
 GO; GO:0008047; F:enzyme activator activity; TAS.
 GO; GO:0005478; F:intracellular transporter activity; TAS.
 GO; GO:0006886; F:intracellular protein transport; TAS.
 InterPro; IPR000379; Ser esters site.
 InterPro; IPR001563; Serine carboxypeptidase.
 Pfam; PF00450; serine carboxypeptidase; 1.
 PRINTS; PR00724; CR50XPTASEC.
 ProDom; PD001189; Serine carboxypeptidase; 2.
 PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
 PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal; Lysosome; 3D-structure; Disease mutation.
 SIGNAL 1 28
 CHAIN 29 480
 CHAIN 29 326
 CHAIN 327 480
 ACT SITE 178 178
 ACT SITE 400 400
 ACT SITE 457 457
 ACT SITE 457 457
 DISULFID 88 362
 DISULFID 240 256
 DISULFID 241 246
 DISULFID 281 331
 CARBOHYD 145 145
 CARBOHYD 333 333
 VARIANT 49 49
 VARIANT 65 65
 VARIANT 90 90
 VARIANT 249 249
 VARIANT 395 395
 VARIANT 440 440
 MUTAGEN 178
 MUTAGEN 457
 CONFLICT 19 19
 CONFLICT 56 56
 HELIX 31 33
 STRAND 34 34
 TURN 39 40
 STRAND 49 55
 TURN 58 59
 STRAND 60 67
 TURN 73 75
 STRAND 78 82
 TURN 85 87
 STRAND 88 88

Query Match 1.4%; Score 8; DB 1; Length 480;
 Best Local Similarity 100.0%; Pred. No. 6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LWLNGGPG 107
 DB 80 LWLNGGPG 87
 RESULT 10
 CBP1_HORVU
 ID CBP1_HORVU STANDARD; PRT; 499 AA.
 AC P07519; P07520;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase C) (CP-MI).
 DE C) (CP-MI).
 GN Hordeum vulgare (Barley).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aleurone;
 RA Rocher A., Lok F., Cameron-Mills V., von Wettstein D.;
 RL Submitted (DSC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 88-499 FROM N.A.
 RX MEDLINE=88298749; PubMed=3403516;
 RA Doan N.P., Fincher G.B.;
 RT "The A- and B-chains of carboxypeptidase I from germinated barley
 RL Originate from a single precursor polypeptide.";
 RN J. Biol. Chem. 263:11106-11110(1988).
 RN [3]
 RP SEQUENCE OF 31-296 AND 352-499.
 RA Soerensen S.B., Breddam K., Svendsen I.;
 RT "Primary structure of carboxypeptidase I from malted barley.";
 RL Carlsberg Res. Commun. 51:475-485(1986).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF SMALL PEPTIDES (2-5 RESIDUES) OR IN THE DEGRADATION OF STORAGE PROTEINS IN THE EMBRYO.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a broad specificity.
 CC -1- SUBUNIT: CARBOXYPEPTIDASE I IS A DIMER, WHERE EACH MONOMER IS COMPOSED OF TWO CHAINS LINKED BY DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.
 CC -1- DEVELOPMENTAL STAGE: AFTER ONE DAY OF GERMINATION, MAINLY FOUND IN THE SCUTELLUM OF THE DEVELOPING GRAIN; BARELY DETECTABLE AFTER FOUR DAYS, AND ABSENT FROM THE MATURE GRAIN. A LOWER LEVEL OF EXPRESSION IS SEEN IN THE ALEURONE BOTH DURING DEVELOPMENT AND GERMINATION.
 CC -1- PTM: THREE DISULFIDE BONDS ARE PRESENT.
 CC -1- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING ENZYME MATURATION
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

PRINTS: PR00724; CRBOXYPTASEC.
 DR PRODM; P0001189; Serine carboxypept; 2.
 DR PROSITE; P000131; CARBOXYPEPT SER SER; 1.
 DR PROSITE; P000560; CARBOXYPEPT SER HIS; 1.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 296
 FT PROPEP 297 351
 FT CHAIN 352 439
 FT ACT SITE 188 188
 FT ACT SITE 423 423
 FT ACT SITE 476 476
 FT CARBOHYD 148 148
 FT CARBOHYD 262 262
 FT CARBOHYD 407 407
 FT SITE 497 499
 FT CONFLICT 102 102
 SQ SEQUENCE 499 AA; 54096 MW; 9C6674B14D9DB9BF CRC64;
 H -> P (IN REF 3).
 Query Match 1.4%; Score 8; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 6.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LWMNGGPG 107
 DB 84 LWMNGGPG 91
 RESULT 11
 CBP1_ORYSA STANDARD; PRT; 510 AA.
 AC P37890;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase C).
 DE GN CBP1.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Yukihikari.
 RX MEDLINE=94213891; PubMed=8161571;
 RA Washio K., Ishikawa K.;
 RT "Cloning and sequencing of the gene for type I carboxypeptidase in rice."
 RL Biochim. Biophys. Acta 1199:311-314(1994).
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a broad specificity.
 CC -1- PPM: THREE DISULFIDE BONDS ARE PRESENT (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D17586; BAA04510.1; -
 CC PIR; S43516; S43516.
 CC HSSP; P08819; 1WHT.
 CC MEROPS; S10.004; -
 CC Gramene; P37890; -
 CC InterPro; IPR000379; Ser estrs site.
 CC InterPro; IPR001563; Serine carboxypept.
 CC Pfam; PF00450; serine carboxypept; 1.
 CC PRINTS; PR00724; CRBOXYPTASEC.

DR PRODM; P0001189; Serine carboxypept; 2.
 DR PROSITE; P000131; CARBOXYPEPT SER SER; 1.
 DR PROSITE; P000560; CARBOXYPEPT SER HIS; 1.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 36
 FT PROPEP 37 510
 FT ACT SITE 194 194
 FT ACT SITE 434 434
 FT ACT SITE 487 487
 FT CARBOHYD 154 154
 FT CARBOHYD 268 268
 FT CARBOHYD 418 418
 FT SITE 508 510
 SQ SEQUENCE 510 AA; 55709 MW; 28896247FA1371CF CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 6.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LWMNGGPG 107
 DB 90 LWMNGGPG 97
 RESULT 12
 CP23_HORVU STANDARD; PRT; 516 AA.
 ID CP23_HORVU
 AC P52711;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine carboxypeptidase II-3 precursor (EC 3.4.16.6) (CP-MII.3).
 DE GN CXP;2-3.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Alexis; TISSUE=Grain;
 RX MEDLINE=9436715; PubMed=7520177;
 RA Dal Degan F., Rocher A., Cameron-Mills V., von Wettstein D.;
 RT "The expression of serine carboxypeptidases during maturation and germination of the barley grain."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213(1994).
 CC -1- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine or lysine residue.
 CC -1- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. ALSO FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.
 CC -1- PPM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING ENZYME MATURATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X78877; CAA55478.1; -
 CC PIR; S44191; S44191.
 CC HSSP; P08819; 1WHT.
 CC MEROPS; S10.005; -
 CC InterPro; IPR000379; Ser estrs site.
 CC InterPro; IPR001563; Serine carboxypept.
 CC Pfam; PF00450; serine carboxypept; 1.
 CC PRINTS; PR00724; CRBOXYPTASEC.

DR ProDom; P0001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 KW Multigene family; Carboxypeptidase; Glycoprotein; Zymogen; Signal;
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 77 POTENTIAL.
 FT CHAIN 78 341 SERINE CARBOXYPEPTIDASE II-3, CHAIN A.
 FT PROPEP 342 352 LINKER PEPTIDE (BY SIMILARITY)
 FT CHAIN 353 516 SERINE CARBOXYPEPTIDASE II-3, CHAIN B.
 FT ACT SITE 236 236 BY SIMILARITY.
 FT ACT SITE 427 427 BY SIMILARITY.
 FT ACT SITE 484 484 BY SIMILARITY.
 FT DISULFID 143 391 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 300 315 BY SIMILARITY.
 FT DISULFID 339 359 BY SIMILARITY.
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 516 AA; 55913 MW; D41AALC56CF8D355 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 516;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LNLNGGPG 107
 Db 135 LNLNGGPG 142
 RESULT 13
 CBPY_PICPA STANDARD; PRT; 523 AA.
 AC P52710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
 GN PRC1.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 108-142.
 RC STRAIN=GTSL15;
 RX MEDLINE=96381245; PubMed=8789258;
 RA Chi H., Ohtani W., Okazaki N., Furuhata N., Ohmura T.;
 RT "Cloning and characterization of the Pichia pastoris PRC1 gene
 encoding carboxypeptidase Y";
 RL Yeast 12:31-40(1996).
 CC -!- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.
 CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 broad specificity.
 CC -!- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X87987; CAA61240.1; -.
 DR PIR; S61713; S61713.
 DR HSSP; P00729; 1YSC.
 DR MEROPS; S10.001; -.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR InterPro; IPR001563; Serine_carbpept.
 DR Pfam; PF00450; serine_carbpept; 1.

DR PRINTS; PR00724; CRBOXYPTASRC.
 DR ProDom; P0001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 107
 FT CHAIN 108 523 CARBOXYPEPTIDASE Y.
 FT ACT SITE 249 249 BY SIMILARITY.
 FT ACT SITE 441 441 BY SIMILARITY.
 FT ACT SITE 498 498 BY SIMILARITY.
 FT DISULFID 162 401 BY SIMILARITY.
 FT DISULFID 296 310 BY SIMILARITY.
 FT DISULFID 320 343 BY SIMILARITY.
 FT DISULFID 327 336 BY SIMILARITY.
 FT DISULFID 365 371 BY SIMILARITY.
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 523 AA; 59447 MW; CA6093BDE2E9D178 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 523;
 Best Local Similarity 100.0%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LNLNGGPG 107
 Db 154 LNLNGGPG 161
 RESULT 14
 CBPY_YEAST STANDARD; PRT; 532 AA.
 AC P00729;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
 GN PRC1 OR YMR297W.
 OS Saccharomycetes cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87131100; PubMed=3028649;
 RA Valls L.A., Hunter C.P., Rothman J.H., Stevens T.H.;
 RT "Protein sorting in yeast: the localization determinant of yeast
 vacuolar carboxypeptidase Y resides in the propeptide";
 RL Cell 48:887-897(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajadream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 XIII";
 RL Nature 387:90-93(1997).
 RN [3]
 RP SEQUENCE OF 112-532.
 RA Svendsen I., Martin B.M., Viswanatha T., Johansen J.T.;
 RT "Amino acid sequence of carboxypeptidase Y. II. Peptides from
 enzymatic cleavages";
 RL Carlsberg Res. Commun. 47:15-27(1982).
 RN [4]
 RP REVISIONS, AND ACTIVE SITE SER-257.
 RA Breddam K., Svendsen I.;
 RT "Identification of methionyl and cysteinyl residues in the substrate
 binding site of carboxypeptidase Y";

FT TURN 467 468
FT HELIX 469 474
FT STRAND 478 481
FT TURN 483 485
FT STRAND 488 492
FT STRAND 494 495
FT TURN 496 497
FT STRAND 498 503
FT STRAND 504 505
FT HELIX 510 513
FT HELIX 515 526
FT TURN 527 529
SQ SEQUENCE 532 AA; 53802 MW; 7227F3489CBDD952 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 6.6; 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 100 LWLNGGPG 107
Db 159 LWLNGGPG 166

RESULT 15

CBEPY_CANAL STANDARD; PRT; 542 AA.
AC P30574;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
GN CPY1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_taxid=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93051356; PubMed=1427093;
RA Mukhtar M., Logan D.A., Kauffer N.F.;
RT "The carboxypeptidase Y-encoding gene from Candida albicans and its
transcription during yeast-to-hyphae conversion.";
RL Gene 121:173-177(1992).
CC -!- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
CC -!- INDUCTION: TRANSIENTLY DOWN-REGULATED DURING THE EARLY EVENTS OF
YEAST TO HYPAE CONVERSION.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95182; AAA34326.2; -.
DR HSSP; P00729; 1CPY.
DR MEROPS; S10.001; -.
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 21
FT PROPEP 22 127 POTENTIAL.
FT CHAIN 128 542 CARBOXYPEPTIDASE Y.

112 Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:02:12 ; Search time 21 Seconds
(without alignments)
2541.601 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 555

Sequence: 1 MRGYEFLSVLPVAASWALP.....HTQSSVPLPTATSMSSVGMA 555

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 76: *
1: piri: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	1.8	423	2 S38953	carboxypeptidase D
2	10	1.8	460	1 S51516	serine-type carbox
3	9	1.6	405	2 S70150	tniQ protein - Xan
4	9	1.6	405	2 T08521	tniQ protein homol
5	9	1.6	468	2 D84503	probable serine ca
6	9	1.6	479	2 G96501	probable serine ca
7	9	1.6	495	2 S32179	tniQ protein homol
8	9	1.6	508	1 S46008	probable carboxype
9	9	1.6	2105	2 T18968	probable serine-type
10	8	1.4	36	2 S78239	ycf32 protein - Od
11	8	1.4	166	2 E71040	hypothetical prote
12	8	1.4	264	2 T49055	hypothetical prote
13	8	1.4	270	2 B75259	serine carboxypept
14	8	1.4	423	1 A29639	hypothetical prote
15	8	1.4	425	2 F85360	carboxypeptidase D
16	8	1.4	425	2 E84631	SERINE CARBOXYPEPT
17	8	1.4	447	2 G84772	probable serine ca
18	8	1.4	452	2 H84772	probable serine ca
19	8	1.4	456	2 H86406	probable serine ca
20	8	1.4	456	2 A85139	probable serine ca
21	8	1.4	458	2 F84746	hypothetical prote
22	8	1.4	465	2 G86244	probable serine ca
23	8	1.4	465	2 B85358	Serine carboxypept
24	8	1.4	470	2 T16606	SERINE CARBOXYPEPT
25	8	1.4	470	2 B96637	probable serine ca
26	8	1.4	474	2 A35732	hypothetical prote
27	8	1.4	476	1 T05701	protective protein
28	8	1.4	480	2 A31589	carboxypeptidase D
29	8	1.4	480	2 T50511	serine-type carbox

30	8	1.4	499	1 CPBHS	carboxypeptidase C
31	8	1.4	501	2 T49081	serine-type carbox
32	8	1.4	502	2 T49188	serin carboxypepti
33	8	1.4	510	1 S43516	carboxypeptidase C
34	8	1.4	510	2 T39601	serine carboxypept
35	8	1.4	512	2 T33463	probable serine ca
36	8	1.4	516	2 S44191	carboxypeptidase D
37	8	1.4	523	1 S61713	carboxypeptidase C
38	8	1.4	532	1 CPBYI	carboxypeptidase C
39	8	1.4	542	1 JC1380	carboxypeptidase C
40	8	1.4	552	2 JC7666	serine-type carbox
41	8	1.4	574	2 T16230	hypothetical prote
42	8	1.4	1002	2 T43236	carboxypeptidase C
43	8	1.4	2338	2 T25810	hypothetical prote
44	7	1.3	25	2 S78343	hypothetical prote
45	7	1.3	62	2 T21953	hypothetical prote

ALIGNMENTS

RESULT 1

S38953

carboxypeptidase D (EC 3.4.16.6) - Penicillium janthinellum

N:Alternate names: carboxypeptidase S1

C:Species: Penicillium janthinellum

C>Date: 18-Feb-1994 #sequence_revision 19-Apr-1996 #text_change 31-Mar-2000

C/Accession: S38953

R:Svendsen, I.; Hofmann, T.; Endrizzi, J.; Remington, S.J.; Breddam, K.

FEBS Lett. 333, 39-43, 1993

A>Title: The primary structure of carboxypeptidase S1 from Penicillium janthinellum.

A/Reference number: S38953; MUID:94039747; PMID:8224168

A/Accession: S38953

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-423 <SVE>

C:Superfamily: serine carboxypeptidase

C/Keywords: hydrolase; serine carboxypeptidase

Query Match

Best Local Similarity 1.8%; Score 10; DB 2; Length 423;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TSYGGHYGP 236

Db 141 TSYGGHYGP 150

RESULT 2

S51516

serine-type carboxypeptidase (EC 3.4.16.-) Z precursor - Absidia zychae

C:Species: Absidia zychae

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Jun-2000

C/Accession: S51516; S78013; S78014

R:Lee, B.R.; Takeuchi, M.; Kobayashi, Y.

Curr. Genet. 27, 159-165, 1995

A>Title: Molecular cloning and sequence analysis of the scp2 gene encoding the serine ca

A/Reference number: S51516; MUID:95308538; PMID:7788719

A/Accession: S51516

A/Molecule type: DNA

A/Residues: 1-460 <LEE>

A/Cross-references: EMBL:D16519; NID:G556466; PIDN:BAA03966.1; PID:g995456

A/Experimental source: strain NRIC 1199

A/Accession: S78013

A/Molecule type: protein

A/Residues: 52-62;90-99;367-381 <LEE>

A/Accession: S78014

A/Molecule type: mRNA

A/Residues: 18-460 <LES>

C:Genetics:

A:Gene: spc2

A/Introns: 64/3; 95/1; 107/2; 138/3; 234/2; 246/1; 270/1; 343/3; 353/2; 384/3; 437/3

C:Superfamily: serine carboxypeptidase

C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-51/Domain: propeptide #status predicted <PRE>
F;52-460/Product: serine-type carboxypeptidase #status experimental <MAT>
F;40,46,119/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;186/Active site: Ser #status predicted

Query Match 1.8%; Score 10; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 486 RYERAGHEVP 495
Db 430 RYERAGHEVP 439
|||||

RESULT 3
S70150
tniQ protein - Xanthomonas sp.
C;Species: Xanthomonas sp.
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 11-Jan-2000
C;Accession: S70150
R;Kholodii, G.Y.; Windlin, S.Z.; Bass, I.A.; Yurieva, O.V.; Minakhina, S.V.; Nikiforov, M.I. Microbiol. 17, 1189-1200, 1995
A;Title: Four genes, two ends, and a res region are involved in transposition of Tn5053:
A;Reference number: S70140; MUID:96130850; PMID:8594337
A;Accession: S70150
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-405 <XHO>
A;Cross-references: EMBL:L40585; NID:g710572; PIDN:AA98331.1; PID:g710584
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C;Genetics:
A;Gene: tniQ
A;Start codon: GTG
A;Superfamily: Xanthomonas tniQ protein
C;Keywords: transposition

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARHPETA 96
Db 341 EARHPETA 349
|||||

RESULT 4
T08521
tniQ protein homolog - Enterobacter aerogenes plasmid R751
C;Species: Enterobacter aerogenes
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: T08521
R;Inorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
Plasmid 36, 95-111, 1996
A;Title: Conservation of the genetic switch between replication and transfer genes of In
A;Reference number: Z16434; MUID:97118926; PMID:8954881
A;Accession: T08521
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-405 <THO>
A;Cross-references: EMBL:U67194; NID:g1572520; PIDN:AA64465.1; PID:g1572570
C;Genetics:
A;Genome: plasmid R751
C;Superfamily: Xanthomonas tniQ protein
C;Keywords: transposition

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARHPETA 96
Db 341 EARHPETA 349
|||||

Db 341 EARHPETA 349

RESULT 5

D84503

probable serine carboxypeptidase II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: D84503
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, I.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84503
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-468 <STO>
C;Cross-references: GB:AE002093; NID:g4733981; PIDN:AA28662.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g12480
A;Map position: 2
C;Superfamily: serine carboxypeptidase

Query Match 1.6%; Score 9; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWNGGPG 107

Db 77 TLWNGGPG 85

RESULT 6

G96501

probable serine carboxypeptidases [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: G96501
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96501
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-479 <STO>
A;Cross-references: GB:AE005173; NID:g7523661; PIDN:AAF63101.1; GSPDB:GN00141
C;Genetics:
A;Gene: F28H19.5
A;Map position: 1
C;Superfamily: serine carboxypeptidase

Query Match 1.6%; Score 9; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWNGGPG 107

Db 83 TLWNGGPG 91

RESULT 7

S32179

tniQ protein homolog - Klebsiella pneumoniae

C;Species: Klebsiella pneumoniae
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 C;Accession: S32179
 R;Radstroem, P.; Sundstrom, L.; Swedberg, G.; Flensburg, J.; Skold, O.
 submitted to the EMBL Data Library, March 1993
 A;Description: The integrons of R751 and fn21 are transposons related to the retroelement
 A;Reference number: S32177
 A;Accession: S32179
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-495 <RAD>
 A;Cross-references: EMBL:X72585; NID:g288626; PIDN:CAA51177.1; PID:g288629
 C;Superfamily: Xanthomonas tniQ protein
 C;Keywords: transposition

Query Match 1.6%; Score 9; DB 2; Length 495;
 Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

Qy 88 EARNPETA 96
 |||||
 Db 431 EARNPETA 439

RESULT 8

S46008
 probable carboxypeptidase C (EC 3.4.16.5) YBR139w - Yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein YBR1015; hypothetical protein YBR139w
 C;Species: Saccharomyces cerevisiae
 C;Date: 26-Aug-1994 #sequence_revision 05-Jan-1996 #text_change 19-Apr-2002
 C;Accession: S46008; S46581
 R;Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
 submitted to the Protein Sequence Database, August 1994
 A;Reference number: S45995
 A;Accession: S46008
 A;Molecule type: DNA
 A;Residues: 1-508 <BEC>
 A;Cross-references: EMBL:Z36008; NID:g536435; PIDN:CAA85097.1; PID:g536436; GSPDB:GN0000
 A;Experimental source: strain S288C
 R;Becam, A.M.; Cullin, C.; Grzybowska, E.; Lacroite, F.; Nasr, F.; Ozier-Kalogeropoulos,
 Yeast 10(Suppl.A), S1-S11, 1994
 A;Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 complete
 A;Reference number: S46569; MUID:94378717; EMID:8091856
 A;Accession: S46581
 A;Molecule type: DNA
 A;Residues: 1-508 <BEC>
 A;Cross-references: EMBL:X75891; NID:g496856; PIDN:CAA53497.1; PID:g496869
 A;Experimental source: strain S288C
 C;Genetics:

A;Cross-references: SGD:S0000343

A;Map position: 2R

A;Note: MIPS:YBR139w

C;Superfamily: serine carboxypeptidase

C;Keywords: hydrolase; serine carboxypeptidase; transmembrane protein

F;6-22/Domain; transmembrane status predicted <TM>

F;219,415,474/Active site: Ser, Asp, His #status predicted

Query Match 1.6%; Score 9; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

Qy 147 QPLGVGFYSY 155
 |||||
 Db 171 QPLGVGFYSY 179

RESULT 9

T18968
 probable serine-type carboxypeptidase (EC 3.4.16.-) Y16B4A.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
 C;Accession: T18968; T23145; T26477
 R;Thomas, K.

submitted to the EMBL Data Library, March 1996
 A;Reference number: Z19053
 A;Accession: T18968
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2105 <WIL>
 A;Cross-references: EMBL:Z70203; PIDN:CAA94110.1; GSPDB:GN000028; CESP:Y16B4A.2
 R;Lloyd, C.
 submitted to the EMBL Data Library, March 1997
 A;Reference number: Z19697
 A;Accession: T23145
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2105 <WI2>
 A;Cross-references: EMBL:Z93339; PIDN:CAB07544.1; GSPDB:GN000028; CESP:Y16B4A.2
 A;Experimental source: clone H40D05
 R;Wallis, J.
 submitted to the EMBL Data Library, June 1998
 A;Reference number: Z20220
 A;Accession: T26477
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2105 <WI3>
 A;Cross-references: EMBL:AL023825; PIDN:CAA19443.1; GSPDB:GN000028; CESP:Y16B4A.2
 A;Experimental source: clone Y16B4A
 C;Comment: This protein has 4 repeats of sequence homologous to serine carboxypeptidase.
 C;Genetics:
 A;Gene: CESP:Y16B4A.2
 A;Map position: X
 A;Introns: 43/3; 67/2; 179/1; 208/3; 239/2; 281/2; 296/3; 346/2; 444/3; 503/3; 588/3; 68
 1664/3; 1752/2; 1794/3; 1834/3; 1882/2; 1905/2; 1942/3; 1982/3
 C;Keywords: duplication; hydrolase; serine carboxypeptidase

Query Match 1.6%; Score 9; DB 2; Length 2105;
 Best Local Similarity 100.0%; Pred. No. 6; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

Qy 100 LWLNGPGS 108
 |||||
 Db 549 LWLNGPGS 557

RESULT 10

S78239

Ycf32 protein - Odontella sinensis chloroplast

C;Species: chloroplast Odontella sinensis

C;Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000

C;Accession: S78239; S78293

R;Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A;Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis

A;Reference number: S78238

A;Accession: S78239

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-36 <KOW>

A;Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91612.1; PID:g1185129

A;Genetics: 5GEN

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

A;Note: 5' gene

A;Accession: S78293

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-36 <KOF>

A;Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91666.1; PID:g1185183

A;Genetics: 3GEN

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

A;Note: 3' gene (c)

C;Genetics: <GEN1>

A;Gene: ycf32

A;Genome: chloroplast

A;Note: gene located on inverted repeat Ira

C;Genetics: <GEN2>
A;Gene: ycf32,
A;Genome: chloroplast
A;Note: gene located on inverted repeat IRB
C;Superfamily: hypothetical protein ycf32
C;Keywords: chloroplast

Query Match 1.4%; Score 8; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

Qy 12 LVAASWAL 19
Db 13 LVAASWAL 20
|||||

RESULT 11
E71040
hypothetical protein PH1613 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C;Accession: E71040
R;Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: E71040
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-166 <RAW>
A;Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30725.1; PID:g3258042
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1613

Query Match 1.4%; Score 8; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

Qy 465 PLKVGVE 472
Db 43 PLKVGVE 50
|||||

RESULT 12
T49055
serine carboxypeptidase-like protein - Arabidopsis thaliana
N;Alternate names: protein TSP19.190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49055
R;Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoorge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25014
A;Accession: T49055
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-264 <BEN>
A;Cross-references: EMBL:AL163972; GSPDB:GN00061; ATSP:TSP19.190
A;Experimental source: cultivar Columbia; BAC clone TSP19
C;Genetics:
A;Gene: ATSP:TSP19.190
A;Map position: 3
A;Introns: 131/1; 166/1

Query Match 1.4%; Score 8; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

Qy 100 LWLNGGPG 107
|||||

Db 126 LWLNGGPG 133

RESULT 13

B75259
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: B75259
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Praser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: B75259
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <WHL>
A;Cross-references: GB:AE002085; GB:AE000513; NID:g6460383; PIDN:AAF12107.1; PID:g6460383
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2564
A;Map position: 1

Query Match 1.4%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

Qy 195 LAEEAAWE 202
Db 166 LAEEAAWE 173
|||||

RESULT 14
A29639
carboxypeptidase D (EC 3.4.16.6) - wheat
N;Alternate names: serine-type carboxypeptidase II
C;Species: Triticum aestivum (common wheat)
C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C;Accession: A29639; C29639; B29639
R;Breddam, K.; Sorensen, S.B.; Svendsen, I.
Carlsberg Res. Commun. 52, 297-311, 1987
A;Title: Primary structure and enzymatic properties of carboxypeptidase II from wheat
A;Reference number: A90771
A;Accession: A29639
A;Molecule type: protein
A;Residues: 1-263 <BRE>
A;Note: A chain
A;Accession: C29639
A;Molecule type: protein
A;Residues: 4-263 <BR2>
A;Note: A' chain
A;Accession: B29639
A;Molecule type: protein
A;Residues: 264-423 <BR3>
A;Note: B chain
C;Complex: homodimer of two mature products, each of which is cleaved to contain an A-
C;Superfamily: serine carboxypeptidase
C;Keywords: disulfide bond; glycoprotein; homodimer; hydrolase; serine carboxypeptidase
F;1-263,264-423/Product: carboxypeptidase D #status experimental <MAT>
F;1-263/Domain: A chain #status experimental <CHA>
F;4-263/Domain: A chain #status experimental <CHAI>
F;264-423/Domain: B chain #status experimental <CHB>
F;116,127,259,291,297,421/Binding site: carboxylate (Asn) (covalent) #status experime
F;158,340,392/Active site: Ser, Asp, His #status predicted

Query Match 1.4%; Score 8; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

Qy 100 LWLNGGPG 107
|||||

Db 57 LWLNGGPG 64

RESULT 15

F85360

SERINE CARBOXYPEPTIDASE II-like protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001

C;Accession: F85360

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: F85360

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-425 <STO>

A;Cross-references: GB:NC_001268; NID:97269982; PIDN:CAB79799.1; GSPDB:GN00140

C;Genetics:

A;Gene: At4g30810

A;Map position: 4

C;Superfamily: serine carboxypeptidase

Query Match

Best Local Similarity 1.4%; Score 8; DB 2; Length 425;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107

|||||||

Db 85 LWLNGGPG 92

Search completed: November 21, 2003, 18:05:49

Job time : 22 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:03:57 ; Search time 36 Seconds
(without alignments)

2814.459 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 555

Sequence: 1 MRGYEFLSVLPVAASWALP.....HTQSSVPLPTATSSVCGMA 555

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.4	209	12	US-10-294-039-4
2	8	1.4	406	12	US-10-194-919-35
3	8	1.4	411	9	US-09-901-252-16
4	8	1.4	421	9	US-09-420-785A-4
5	8	1.4	421	9	US-09-901-252-15
6	8	1.4	480	10	US-09-736-457-336
7	8	1.4	480	10	US-09-902-941-336
8	8	1.4	480	10	US-09-849-626-336
9	8	1.4	480	11	US-09-476-300-336
10	8	1.4	480	12	US-10-113-872-336
11	8	1.4	480	14	US-10-084-018-8
12	8	1.4	480	15	US-10-017-754-336
13	8	1.4	492	9	US-09-925-302-774
14	8	1.4	492	11	US-09-796-753-42
15	8	1.4	2991	12	US-09-970-944-22

16	8	1.4	4349	12	US-10-174-677-76	Sequence 76, Appl
17	8	1.4	4349	12	US-09-970-944-4	Sequence 4, Appl
18	8	1.4	4349	12	US-09-970-944-18	Sequence 18, Appl
19	8	1.4	4349	12	US-09-970-944-19	Sequence 19, Appl
20	8	1.4	4349	15	US-10-160-758-15	Sequence 15, Appl
21	7	1.3	16	12	US-10-158-825-96	Sequence 96, Appl
22	7	1.3	16	15	US-10-158-847-96	Sequence 96, Appl
23	7	1.3	17	15	US-10-062-710-131	Sequence 131, Appl
24	7	1.3	34	9	US-09-776-490-24	Sequence 24, Appl
25	7	1.3	34	9	US-09-776-491-24	Sequence 24, Appl
26	7	1.3	70	9	US-09-867-550-112	Sequence 112, Appl
27	7	1.3	71	9	US-09-864-761-40430	Sequence 40430, A
28	7	1.3	72	10	US-09-854-512-1	Sequence 1, Appl
29	7	1.3	77	9	US-09-864-761-41457	Sequence 41457, A
30	7	1.3	85	10	US-09-751-100B-54	Sequence 54, Appl
31	7	1.3	85	10	US-09-751-100B-55	Sequence 55, Appl
32	7	1.3	85	10	US-09-751-100B-56	Sequence 56, Appl
33	7	1.3	85	10	US-09-751-100B-57	Sequence 57, Appl
34	7	1.3	85	10	US-09-751-100B-58	Sequence 58, Appl
35	7	1.3	85	14	US-10-092-243A-16	Sequence 16, Appl
36	7	1.3	91	12	US-09-933-767-584	Sequence 584, App
37	7	1.3	91	15	US-10-023-282-584	Sequence 8, Appl
38	7	1.3	163	9	US-09-735-487-8	Sequence 8, Appl
39	7	1.3	163	12	US-10-342-188-8	Sequence 8, Appl
40	7	1.3	163	15	US-10-156-761-12744	Sequence 12744, A
41	7	1.3	193	10	US-09-738-626-3874	Sequence 3874, Ap
42	7	1.3	225	10	US-09-908-805B-8	Sequence 8, Appl
43	7	1.3	229	10	US-09-738-626-5578	Sequence 5578, Ap
44	7	1.3	271	9	US-09-901-884-4	Sequence 4, Appl
45	7	1.3	317	12	US-10-333-631-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-294-039-4
; Sequence 4, Application US/10294039
; Publication No. US20030134814A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph M.
; TITLE OF INVENTION: METHOD OF USING 18080, A HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDE FAMILY MEMBER
; FILE REFERENCE: MP101-230PIRM
; CURRENT APPLICATION NUMBER: US/10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/338,587
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(209)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-294-039-4

Query Match 1.4%; Score 8; DB 12; Length 209;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGPG 107

Db 33 LWLNGPG 40

RESULT 2

US-10-194-919-35
; Sequence 35, Application US/10194919
; Publication No. US20030145350A1

```

; GENERAL INFORMATION:
; APPLICANT: GVS Gesellschaft fuer Erwerb und Verwertung von Sc
; APPLICANT: Spener, Friedrich
; APPLICANT: Abbadi, Amine
; APPLICANT: Brummel, Monika
; TITLE OF INVENTION: METHOD FOR INCREASING THE CONTENT OF FATTY ACIDS IN PLANTS AND
; FILE REFERENCE: 1617.018US1
; CURRENT APPLICATION NUMBER: US/10/194,919
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: PCT/EP01/00289
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DE 100 00 978.6
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-194-919-35

Query Match      1.4%; Score 8; DB 12; Length 406;
Best Local Similarity 100.0%; Pred. No. 28;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      193  TNLAEAA 200
Db      121  TNLAEAA 128
|||||

RESULT 3
US-09-901-252-16
; Sequence 16, Application US/09901252
; Patent No. US20020026658A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: Genes Encoding Sinapoylglucose:Malate Sinapoyltransferase and Met
; FILE REFERENCE: N1422-004
; CURRENT APPLICATION NUMBER: US/09/901,252
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216593
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Triticales
US-09-901-252-16

Query Match      1.4%; Score 8; DB 9; Length 411;
Best Local Similarity 100.0%; Pred. No. 28;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100  LWLNGGPG 107
Db      57  LWLNGGPG 64
|||||

RESULT 4
US-09-420-785A-4
; Sequence 4, Application US/09420785A
; Patent No. US20010010923A1
; GENERAL INFORMATION:
; APPLICANT: MORTENSEN, UFFE
; APPLICANT: OLESEN, KJELD
; APPLICANT: STENNICKE, HENNING
; APPLICANT: SORENSEN, STEEN B.
; APPLICANT: BREDDAM, KLAUS
; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
; FILE REFERENCE: 089187/0109

```

```

; CURRENT APPLICATION NUMBER: US/09/420,785A
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-420-785A-4

Query Match      1.4%; Score 8; DB 9; Length 421;
Best Local Similarity 100.0%; Pred. No. 29;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100  LWLNGGPG 107
Db      48  LWLNGGPG 55
|||||

RESULT 5
US-09-901-252-15
; Sequence 15, Application US/09901252
; Patent No. US20020026658A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: Genes Encoding Sinapoylglucose:Malate Sinapoyltransferase and Me
; FILE REFERENCE: N1422-004
; CURRENT APPLICATION NUMBER: US/09/901,252
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216593
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccaromyces
US-09-901-252-15

Query Match      1.4%; Score 8; DB 9; Length 421;
Best Local Similarity 100.0%; Pred. No. 29;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100  LWLNGGPG 107
Db      48  LWLNGGPG 55
|||||

RESULT 6
US-09-736-457-336
; Sequence 336, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Alijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-336

Query Match      1.4%; Score 8; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 LWLNGGPG 107
Db      80 LWLNGGPG 87

RESULT 7
US-09-902-941-336
; Sequence 336, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-336

Query Match      1.4%; Score 8; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 LWLNGGPG 107
Db      80 LWLNGGPG 87

RESULT 8
US-09-849-626-336
; Sequence 336, Application US/09849626
; Publication No. US2002019769A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-336

Query Match      1.4%; Score 8; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 LWLNGGPG 107
Db      80 LWLNGGPG 87

RESULT 9
US-09-476-300-336
; Sequence 336, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-476-300-336

Query Match      1.4%; Score 8; DB 11; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 LWLNGGPG 107
Db      80 LWLNGGPG 87

RESULT 10
US-10-113-872-336
; Sequence 336, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-113-872-336

Query Match      1.4%; Score 8; DB 12; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 LWLNGGPG 107
Db      80 LWLNGGPG 87

```

APPLICANT: Marnerakis, Margarita
 APPLICANT: Carter, Darrick
 APPLICANT: Fanger, Gary R.
 APPLICANT: Vedwick, Thomas S.
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: McNabb, Andria
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.478C18
 CURRENT APPLICATION NUMBER: US/10/017,754
 CURRENT FILING DATE: 2001-10-29
 NUMBER OF SEQ ID NOS: 2004
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 336
 LENGTH: 480
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-017-754-336

Query Match 1.4%; Score 8; DB 15; Length 480;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
 Db 80 LWLNGGPG 87

RESULT 13
 US-09-925-302-774
 Sequence 774, Application US/09925302
 Patent No. US20020044941A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PAL04
 CURRENT APPLICATION NUMBER: US/09/925,302
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05918
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 896
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 774
 LENGTH: 492
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-925-302-774

Query Match 1.4%; Score 8; DB 9; Length 492;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
 Db 92 LWLNGGPG 99

RESULT 14
 US-09-796-753-42
 Sequence 42, Application US/09796753
 Publication No. US20030027998A1
 GENERAL INFORMATION:
 APPLICANT: McCarthy, Sean A.
 TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 FILE REFERENCE: 7853-227-999
 CURRENT APPLICATION NUMBER: US/09/796,753
 CURRENT FILING DATE: 2001-03-01
 PRIOR APPLICATION NUMBER: 09/183,175
 PRIOR FILING DATE: 1998-10-30
 PRIOR APPLICATION NUMBER: 09/223,094
 PRIOR FILING DATE: 1998-12-30

RESULT 11
 US-10-084-018-8
 Sequence 8, Application US/10084018
 Publication No. US20020160499A1
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 Hawkins, Phillip R.
 Hillman, Jennifer L.
 Lai, Preeti
 Goli, Surya K.
 TITLE OF INVENTION: NOVEL HUMAN SERINE
 CARBOXYPEPTIDASE
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/084,018
 FILING DATE: 25-Feb-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,689A
 FILING DATE: <Unknown>
 APPLICATION NUMBER: 08/828,488
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PP-0241 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 480 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 190283
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-10-084-018-8

Query Match 1.4%; Score 8; DB 14; Length 480;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
 Db 80 LWLNGGPG 87

RESULT 12
 US-10-017-754-336
 Sequence 336, Application US/10017754
 Publication No. US20030054363A1
 GENERAL INFORMATION:
 APPLICANT: Henderson, Robert A.
 APPLICANT: Wang, Tongtong
 APPLICANT: Watanabe, Yoshihiro
 APPLICANT: Johnson, Jeffrey C.
 APPLICANT: Retter, Marc W.

; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 42
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-42

Query Match 1.4%; Score 8; DB 11; Length 492;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 93 LWLNGGPG 100

RESULT 15

US-09-970-944-22
; Sequence 22, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Heriman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1e1 Proteins and Nucleic Acids Encoding Same an

; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 2991
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-22

Query Match 1.4%; Score 8; DB 12; Length 2991;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 INPVTGVV 170
Db 1278 INPVTGVV 1285

Search completed: November 21, 2003, 18:07:11
Job time : 37 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:03:02 ; Search time 22 Seconds
(without alignments)
1067.388 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 555

Sequence: 1 MRGYEFLSVLPLVAASWALP.....HTQSSVPLTATSMSSVGMA 555

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	99.8	554	3	US-08-943-714-2
2	35	6.3	35	3	US-08-943-714-5
3	19	3.4	20	3	US-08-943-714-3
4	16	2.9	29	3	US-08-943-714-6
5	10	1.8	423	3	US-08-943-714-9
6	8	1.4	81	4	US-09-439-554-2
7	8	1.4	283	4	US-09-107-532A-5418
8	8	1.4	291	4	US-09-439-554-20
9	8	1.4	421	2	US-08-807-263-4
10	8	1.4	480	2	US-08-828-488-8
11	8	1.4	480	4	US-09-299-689A-8
12	8	1.4	480	4	US-09-702-705-336
13	8	1.4	480	4	US-09-736-457-336
14	8	1.4	491	1	US-09-640-305-4
15	8	1.4	491	1	US-08-360-673-4
16	8	1.4	532	2	US-08-899-324-33
17	8	1.4	532	3	US-08-329-892B-33
18	8	1.4	557	1	US-08-309-341-2
19	8	1.4	557	1	US-08-309-341-4
20	8	1.4	557	1	US-08-608-267-2
21	8	1.4	557	1	US-08-608-267-4
22	8	1.4	557	1	US-08-608-452-2
23	8	1.4	557	1	US-08-608-452-4
24	8	1.4	557	1	US-08-608-224-2
25	8	1.4	557	1	US-08-608-224-4
26	8	1.4	557	1	US-08-967-149-2
27	8	1.4	557	2	US-08-967-149-4

```

28 7 1.3 9 3 US-08-159-339A-528 Sequence 528, App
29 7 1.3 10 3 US-08-159-339A-549 Sequence 549, App
30 7 1.3 19 3 US-08-943-714-4 Sequence 4, Appli
31 7 1.3 34 3 US-08-810-009-24 Sequence 24, Appl
32 7 1.3 55 2 US-08-637-759B-348 Sequence 348, App
33 7 1.3 55 3 US-08-871-355A-348 Sequence 348, App
34 7 1.3 55 4 US-09-201-945-348 Sequence 54, Appl
35 7 1.3 85 3 US-08-894-173-54 Sequence 54, Appl
36 7 1.3 85 3 US-08-894-173-55 Sequence 55, Appl
37 7 1.3 85 3 US-08-894-173-56 Sequence 56, Appl
38 7 1.3 85 3 US-08-894-173-57 Sequence 57, Appl
39 7 1.3 85 3 US-08-894-173-58 Sequence 58, Appl
40 7 1.3 85 3 US-08-398-193-54 Sequence 54, Appl
41 7 1.3 85 3 US-09-398-193-55 Sequence 55, Appl
42 7 1.3 85 3 US-09-398-193-56 Sequence 56, Appl
43 7 1.3 85 3 US-09-398-193-57 Sequence 57, Appl
44 7 1.3 85 3 US-09-398-193-58 Sequence 58, Appl
45 7 1.3 91 4 US-09-205-258-584 Sequence 584, App

```

ALIGNMENTS

```

RESULT 1
US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dambmann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61875780 No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-943-714-2

```

Query Match 99.8%; Score 554; DB 3; Length 554;

FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6187578e
US-08-943-714-5

Query Match 6.3%; Score 35; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 4.4e-29;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 DAIGNVNTQSNNDVYAFQQTGDFVWPNFIEDL 417
Db 1 DAIGNVNTQSNNDVYAFQQTGDFVWPNFIEDL 35

RESULT 3
US-08-943-714-3
Sequence 3, Application US/08943714
Patent No. 6187578
GENERAL INFORMATION:
APPLICANT: Blinkovsky, Alexander
APPLICANT: Berka, Randy
APPLICANT: Rey, Michael
APPLICANT: Golightly, Elizabeth
APPLICANT: Klotz, Alan
APPLICANT: Mathisen, Thomas Erik
APPLICANT: Dammann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

Best Local Similarity 100.0%; Pred. No. 0;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGYEFLSVLPVAAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRYKPGAGV 60
Db 1 MRGYEFLSVLPVAAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRYKPGAGV 60

QY 61 CTTTGVSYSYGVDTSPESHFFPWFEEARHNPETAPITLWNGPGSGDSLGLFEEELGP 120
Db 61 CTTTGVSYSYGVDTSPESHFFPWFEEARHNPETAPITLWNGPGSGDSLGLFEEELGP 120

QY 121 CHVNSTFDYINPHSMNEVSNLLFLSQPLGVGFSDYTDVGSINPVTGVWVSSAGVQG 180
Db 121 CHVNSTFDYINPHSMNEVSNLLFLSQPLGVGFSDYTDVGSINPVTGVWVSSAGVQG 180

QY 181 RYPTIDATLIDTNLAAEAWEILOGLFLSLDSRVQSKDFSLWTESYGGHYGPAPFN 240
Db 181 RYPTIDATLIDTNLAAEAWEILOGLFLSLDSRVQSKDFSLWTESYGGHYGPAPFN 240

QY 241 HVEQNERIANGSVQVNFNSLGIINGIIDEAIAQYPPPEFVANNVTYGIKAVNETVYN 300
Db 241 HVEQNERIANGSVQVNFNSLGIINGIIDEAIAQYPPPEFVANNVTYGIKAVNETVYN 300

QY 301 YMKFANQMPNGCQDLISTCKQNTRTALADYALCAEATNMCRDNVEGPIYAFAGRGYD 360
Db 301 YMKFANQMPNGCQDLISTCKQNTRTALADYALCAEATNMCRDNVEGPIYAFAGRGYD 360

QY 361 HPYDDPTPSYKFKLAKSVMDAIGNVNTQSNNDVYAFQQTGDFVWPNFIEDLEE 420
Db 361 HPYDDPTPSYKFKLAKSVMDAIGNVNTQSNNDVYAFQQTGDFVWPNFIEDLEE 420

QY 421 LALPVRVSLIYGADYICNWFQGVQAVSLAANYQAAQFRSAGYTPLKVNGVEYGETREYG 480
Db 421 LALPVRVSLIYGADYICNWFQGVQAVSLAANYQAAQFRSAGYTPLKVNGVEYGETREYG 480

QY 481 NFSFTRVYAGHVVPYQPIASQLNRTIFGWDIAEGQKKIWFPSYKXTGTATATHTQSS 540
Db 481 NFSFTRVYAGHVVPYQPIASQLNRTIFGWDIAEGQKKIWFPSYKXTGTATATHTQSS 540

QY 541 VPLPTATSMSSVGM 554
Db 541 VPLPTATSMSSVGM 554

RESULT 2
US-08-943-714-5
Sequence 5, Application US/08943714
Patent No. 6187578
GENERAL INFORMATION:
APPLICANT: Blinkovsky, Alexander
APPLICANT: Berka, Randy
APPLICANT: Rey, Michael
APPLICANT: Golightly, Elizabeth
APPLICANT: Klotz, Alan
APPLICANT: Mathisen, Thomas Erik
APPLICANT: Dammann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714


```
; MOLECULE TYPE: No. 6187578e
US-08-943-714-3
Query Match          3.4%; Score 19; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 YGGHYGPAFFNHFYEQNER 248
   |||||
Db 2 YGGHYGPAFFNHFYEQNER 20

RESULT 4
US-08-943-714-6
; Sequence 6, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berkka, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 12
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6187578e
US-08-943-714-6
Query Match          2.9%; Score 16; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 YAFAGRGVYDIRPYD 364
   |||||
Db 11 YAFAGRGVYDIRPYD 26

RESULT 5
US-08-943-714-9
; Sequence 9, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berkka, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6187578e
US-08-943-714-9
Query Match          1.8%; Score 10; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.077;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 TSYGGHYGP 236
   |||||
Db 141 TSYGGHYGP 150

RESULT 6
US-09-439-554-2
; Sequence 2, Application US/09439554
; Patent No. 6479733
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Jan Antoni
; APPLICANT: Odell, Joan T.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: STEROL METABOLISM ENZYMES
; FILE REFERENCE: BB1114 US NA
; CURRENT APPLICATION NUMBER: US/09/439,554
; CURRENT FILING DATE: 1999-11-12
; EARLIER APPLICATION NUMBER: 60/108,351
; EARLIER FILING DATE: 1998-No. 6479733ember-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 81
```

```

; TYPE: PRT
; ORGANISM: Zea mays
US-09-439-554-2

Query Match      1.4%; Score 8; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      207 FLUGLPSL 214
DB      29 FLUGLPSL 36

RESULT 7
US-09-107-532A-5418
; Sequence 5418, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5418:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...283
; SEQUENCE DESCRIPTION: SEQ ID NO: 5418:
US-09-107-532A-5418

Query Match      1.4%; Score 8; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      379 DSVMDAIG 386
DB      187 DSVMDAIG 194
```

```

RESULT 8
US-09-439-554-20
; Sequence 20, Application US/09439554
; Patent No. 6479733
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Jan Antoni
; APPLICANT: Odell, Joan T.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Pamodu, Omelayo O.
; TITLE OF INVENTION: STEROL METABOLISM ENZYMES
; FILE REFERENCE: BB1114 US NA
; CURRENT APPLICATION NUMBER: US/09/439,554
; CURRENT FILING DATE: 1999-11-12
; EARLIER APPLICATION NUMBER: 60/108,351
; EARLIER FILING DATE: 1998-No. 6479733ember-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Zea mays
US-09-439-554-20

Query Match      1.4%; Score 8; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      207 FLUGLPSL 214
DB      61 FLUGLPSL 68

RESULT 9
US-08-807-263-4
; Sequence 4, Application US/08807263C
; Patent No. 5985627
; GENERAL INFORMATION:
; APPLICANT: Mortensen, Uffe
; APPLICANT: Olesen, Kjeld
; APPLICANT: Stennicke, Henning
; APPLICANT: Sorensen, Steen B.
; APPLICANT: Bredam, Klaus
; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
; FILE REFERENCE: 8648.71us01-no4
; CURRENT APPLICATION NUMBER: US/08/807,263C
; CURRENT FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-08-807-263-4

Query Match      1.4%; Score 8; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 LWLNGGPG 107
DB      48 LWLNGGPG 55

RESULT 10
US-08-828-488-8
; Sequence 8, Application US/08828488
; Patent No. 5925521
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
```

```

; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,488
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 190283
; US-08-828-488-8

```

```

Query Match 1.4%; Score 8; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 80 LWLNGGPG 87

```

```

RESULT 11
US-09-299-689A-8
; Sequence 8, Application US/09299689A
; Patent No. 6379913
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,689A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,488
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 190283
; US-09-299-689A-8

```

```

Query Match 1.4%; Score 8; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 80 LWLNGGPG 87

```

```

RESULT 12
US-09-702-705-336
; Sequence 336, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Baugur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-702-705-336

```

```

Query Match 1.4%; Score 8; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 80 LWLNGGPG 87

RESULT 13
US-09-736-457-336

```

```

; Sequence 336, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-736-457-336

Query Match 1.4%; Score 8; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
DB 80 LWLNGGPG 87

RESULT 14
US-09-640-305-4
; Sequence 4, Application US/09640305
; Patent No. RE37447
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/640,305
; FILING DATE: 16-Aug-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 06-FEB-1995
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-640-305-4

Query Match 1.4%; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
DB 113 LWLNGGPG 120

RESULT 15
US-08-360-673-4
; Sequence 4, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-360-673-4

Query Match 1.4%; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107

```

Mon Nov 24 13:41:06 2003

us-09-712-338-2.oli.ra1

Page 7

Db 113 LWLGGPG 120

Search completed: November 21, 2003, 18:06:23
Job time : 23 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:00:11 ; Search time 46 Seconds
(without alignments)
1915.071 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 555

Sequence: 1 MRGYEFLSVLPLVAASWALP.....HTQSSVPLPTATSSVSGMA 555

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq_19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	99.8	554	19 AAW56099	A. oryzae ATCC2038
2	35	6.3	35	19 AAW57040	A. oryzae ATCC2038
3	24	4.3	551	23 ABR38864	A. niger serine ca
4	19	3.4	20	19 AAW56100	A. oryzae ATCC2038
5	16	2.9	29	19 AAW57041	A. oryzae ATCC2038
6	10	1.8	526	23 ABR38817	A. niger serine ca
7	9	1.6	178	22 AAU46801	Propionibacterium
8	9	1.6	454	21 AAG17034	Arabidopsis thalia
9	9	1.6	461	21 AAG17033	Arabidopsis thalia

10	9	1.6	461	21	AAG49759	Arabidopsis thalia
11	9	1.6	461	21	AAG49762	Arabidopsis thalia
12	9	1.6	473	21	AAG17032	Arabidopsis thalia
13	9	1.6	473	21	AAG49758	Arabidopsis thalia
14	9	1.6	473	21	AAG49761	Arabidopsis thalia
15	9	1.6	508	23	ABG33049	S. cerevisiae BAX-
16	9	1.6	554	23	ABR38859	A. niger serine ca
17	9	1.6	623	23	ABR38865	A. niger serine ca
18	8	1.4	81	24	ABU07821	Corn C-4 methyl st
19	8	1.4	93	22	AAO07036	Human polypeptide
20	8	1.4	116	23	ABP03224	Human ORFX protein
21	8	1.4	216	15	ABR72106	Bovine growth horm
22	8	1.4	240	21	AAG16761	Arabidopsis thalia
23	8	1.4	246	21	AAG16760	Arabidopsis thalia
24	8	1.4	291	24	ABU07830	Corn C-4 methyl st
25	8	1.4	303	22	ABG06385	Novel human diagno
26	8	1.4	438	18	AAW15092	Human protective p
27	8	1.4	444	23	ABR54514	Lactococcus lactis
28	8	1.4	450	21	AAG29802	Arabidopsis thalia
29	8	1.4	452	18	AAW15091	Human precursor pr
30	8	1.4	452	21	AAG05734	Arabidopsis thalia
31	8	1.4	452	21	AAG06851	Arabidopsis thalia
32	8	1.4	456	21	AAG05733	Arabidopsis thalia
33	8	1.4	458	21	AAG29801	Arabidopsis thalia
34	8	1.4	470	21	AAG29800	Arabidopsis thalia
35	8	1.4	479	21	AAG23905	Arabidopsis thalia
36	8	1.4	479	21	AAG47178	Arabidopsis thalia
37	8	1.4	480	22	ABR76860	Human lung tumour
38	8	1.4	480	23	AAU85515	Clone #18991 of lu
39	8	1.4	480	24	ABU69487	Human lung cancer
40	8	1.4	480	24	ABU66389	Lung cancer therap
41	8	1.4	486	21	AAG23904	Arabidopsis thalia
42	8	1.4	486	21	AAG47177	Arabidopsis thalia
43	8	1.4	491	15	ABR48059	Sequence of protea
44	8	1.4	492	21	ABR58436	Lung cancer associ
45	8	1.4	492	21	ABR01416	Mouse TANGO 176.

ALIGNMENTS

RESULT 1
AAW56099
ID AAW56099 standard; Protein; 554 AA.
XX
AC AAW56099;

XX
27-AUG-1998 (first entry)

DE A. oryzae ATCC20386 carboxypeptidase I protein.

XX Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous; food industry.

OS Aspergillus oryzae.

PH Key Location/Qualifiers
Peptide 1..18
FT /label= signal

FT Protein 19..554
FT /label= carboxypeptidase I

XX WO9814599-A1.

XX 09-APR-1998.

XX 03-OCT-1997; 97WO-US17977.

XX 27-NOV-1996; 96US-0757534.

XX 04-OCT-1996; 96US-0726880.

PA (NOVO) NOVO NORDISK BIOTECH INC.

PA (NOVO) NOVO-NORDISK AS.

XX Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;
PI Klotz A, Mathisen TE, Rey M;
XX WPI; 1998-240098/21.
DR N-PSDB; AAV28620.
XX Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
PT substrates, useful for improving flavour of foods
XX Claim 1; Fig 3; 82pp; English.
XX This sequence represents carboxypeptidase I from Aspergillus oryzae.
CC This polypeptide has an optimal activity in the range of pH 3.0-7.5 at
CC 25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4, and a
CC residual activity of at least 65.5% after 30 minutes at pH 4.0 and
CC 60 deg. C. It also has the capacity to hydrolyse X from N-CBZ-Ala-X where
CC N-CBZ is N-carboxybenzoyl and X is any amino acid. The carboxypeptidases
CC can be used for obtaining hydrolysates (which can be enriched in free
CC glutamic acid or peptide bound glutamic acid residues) from proteinaceous
CC substrates. The carboxypeptidases can be used in flavour-improving
CC compositions in the food industry. The products can also be used for the
CC production of polypeptides free of carboxypeptidase activity.
XX SQ Sequence 554 AA;
Query Match 99.8%; Score 554; DB 19; Length 554;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRGYRFLSVLPLVAASWALPGSTPASVGRRLPKNTGKTLTANNVIRYKEPGAEGV 60
Db 1 MRGYRFLSVLPLVAASWALPGSTPASVGRRLPKNTGKTLTANNVIRYKEPGAEGV 60
QY 61 CETTRGVKSGVVDTSPESTHFFWFEARHNPTAPITLWNGFGSGSLGLFRELGP 120
Db 61 CETTRGVKSGVVDTSPESTHFFWFEARHNPTAPITLWNGFGSGSLGLFRELGP 120
QY 121 CHVNSFTDDYINPHSNWEVSNLLFLSQPLGVGFSYSDTVGSDINPVTGVVENSFFAGVQG 180
Db 121 CHVNSFTDDYINPHSNWEVSNLLFLSQPLGVGFSYSDTVGSDINPVTGVVENSFFAGVQG 180
QY 181 RYPTIDATLIDTNNLAARAAWEILQGLSLGSLDSRVQSKDFSLWTESYGGHYGPAFN 240
Db 181 RYPTIDATLIDTNNLAARAAWEILQGLSLGSLDSRVQSKDFSLWTESYGGHYGPAFN 240
QY 241 HFEQNERIANGSVNGVQLNFNSLGIINGIIDRAIQAPYYPEFVAVNNTYGIKAVNETVYN 300
Db 241 HFEQNERIANGSVNGVQLNFNSLGIINGIIDRAIQAPYYPEFVAVNNTYGIKAVNETVYN 300
QY 301 YMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDNVEGYPYAFAGRGVYDIR 360
Db 301 YMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDNVEGYPYAFAGRGVYDIR 360
QY 361 HPYDDPTPPSYNKFALKDSVMDAIGVINYTSNDNDVYFAFOQTGDFVWPNFIEDLEEI 420
Db 361 HPYDDPTPPSYNKFALKDSVMDAIGVINYTSNDNDVYFAFOQTGDFVWPNFIEDLEEI 420
QY 421 LALPVRVSLYGDADYICNWFVGQAVSLAANYSOAQRAGYTPLVKVGVEYGETREYG 480
Db 421 LALPVRVSLYGDADYICNWFVGQAVSLAANYSOAQRAGYTPLVKVGVEYGETREYG 480
QY 481 NFSFTRVYVYAGHEVYQPIASLQFNRTIFGMDIAEGQKKIWPSTYKNGTATATHTQSS 540
Db 481 NFSFTRVYVYAGHEVYQPIASLQFNRTIFGMDIAEGQKKIWPSTYKNGTATATHTQSS 540
QY 541 VPLPTATSMSSVGM 554
Db 541 VPLPTATSMSSVGM 554
RESULT 2
AAW57040

ID AAW57040 standard; Protein; 35 AA.
XX AAW57040;
AC AAW57040;
XX 27-AUG-1998 (first entry)
DT 27-AUG-1998 (first entry)
XX A. oryzae ATCC20386 carboxypeptidase I peptide fragment #3.
DE Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
KW food industry.
XX Aspergillus oryzae.
OS Aspergillus oryzae.
XX WO9814599-A1.
EN 09-APR-1998.
XX 03-OCT-1997; 97WO-US17977.
PF 27-NOV-1996; 96US-0757534.
XX 04-OCT-1996; 96US-0726880.
PR (NOVO) NOVO NORDISK BIOTECH INC.
XX (NOVO) NOVO-NORDISK AS.
PA Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;
XX Klotz A, Mathisen TE, Rey M;
PI WPI; 1998-240098/21.
XX Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
PT substrates, useful for improving flavour of foods
XX Example 2; Page 42; 82pp; English.
PS This sequence represents a peptide fragment from carboxypeptidase I
XX isolated from Aspergillus oryzae and generated from cyanogen bromide
CC cleavage. this polypeptide has an optimal activity in the range of pH
CC 3.0-7.5 at 25 deg. C, optimal activity in the range of 55-60 deg. C at
CC pH 4, and a residual activity of at least 65.5% after 30 minutes at pH
CC 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from
CC N-CBZ-Ala-X where N-CBZ is N-carboxybenzoyl and X is any amino acid. The
CC carboxypeptidases can be used for obtaining hydrolysates (which can be
CC enriched in free glutamic acid or peptide bound glutamic acid residues)
CC from proteinaceous substrates. The carboxypeptidases can be used in
CC flavour-improving compositions in the food industry. The products can
CC also be used for the production of polypeptides free of carboxypeptidase
CC activity.
XX SQ Sequence 35 AA;
Query Match 6.3%; Score 35; DB 19; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.1e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 383 DAIGVINYTSNDNDVYFAFOQTGDFVWPNFIEDL 417
Db 1 DAIGVINYTSNDNDVYFAFOQTGDFVWPNFIEDL 35
RESULT 3
ABR38864
ID ABR38864 standard; Protein; 551 AA.
XX ABR38864;
XX 24-APR-2003 (first entry)
DT A. niger serine carboxypeptidase polypeptide #3.
XX Protease; fungal infection; aspergillosis; food; tanning; detergent;
KW protein solubility; viscosity; taste; texture; nutritional value;
KW EC3.4.16.6.

XX OS Aspergillus niger.
 XX PN WO200268623-A2.
 XX PD 06-SEP-2002.
 XX PF 22-FEB-2002; 2002WO-EP01984.
 XX PR 23-FEB-2001; 2001EP-0200657.
 XX PR 23-FEB-2001; 2001EP-0200658.
 XX PR 23-FEB-2001; 2001EP-0200660.
 XX PR 26-FEB-2001; 2001EP-0200706.
 XX PR 26-FEB-2001; 2001EP-0200707.
 XX PR 26-FEB-2001; 2001EP-0200708.
 XX PR 26-FEB-2001; 2001EP-0200719.
 XX PR 28-MAR-2001; 2001EP-0000075.
 XX PR 28-MAR-2001; 2001EP-0000078.
 XX PR 28-MAR-2001; 2001EP-0000080.
 XX PR 28-MAR-2001; 2001EP-0000087.
 XX PR 28-MAR-2001; 2001EP-0000088.
 XX PR 21-MAY-2001; 2001EP-0000156.
 XX PR 21-MAY-2001; 2001EP-0000159.
 XX PR 21-MAY-2001; 2001EP-0000160.
 XX PR 21-MAY-2001; 2001EP-0000162.
 XX PR 21-MAY-2001; 2001EP-0000165.
 XX PR 21-MAY-2001; 2001EP-0000166.
 XX PR 21-MAY-2001; 2001EP-0000168.
 XX PR 21-JUN-2001; 2001EP-0000240.
 XX PR 21-JUN-2001; 2001EP-0000242.
 XX PR 21-JUN-2001; 2001EP-0000244.
 XX PR 21-JUN-2001; 2001EP-0000246.
 XX PR 12-JUL-2001; 2001EP-0000280.
 XX PR 12-JUL-2001; 2001EP-0000285.
 XX PR 30-JUL-2001; 2001EP-0000323.
 XX PR 30-JUL-2001; 2001EP-0000327.
 XX PR 02-AUG-2001; 2001EP-0000341.
 XX PR 02-AUG-2001; 2001EP-0000342.
 XX PR 02-AUG-2001; 2001EP-0000343.
 XX PR 02-AUG-2001; 2001EP-0000344.
 XX PR 09-AUG-2001; 2001EP-0000357.
 XX PR 16-AUG-2001; 2001EP-0000374.
 XX PR 16-AUG-2001; 2001EP-0000377.
 XX PR 20-SEP-2001; 2001EP-0000478.
 XX PR 20-SEP-2001; 2001EP-0000483.
 XX PR 22-OCT-2001; 2001EP-0000552.
 XX PR 22-OCT-2001; 2001EP-0000553.
 XX PR 22-OCT-2001; 2001EP-0000554.
 XX PR 22-OCT-2001; 2001EP-0000556.
 XX PR 22-OCT-2001; 2001EP-0000557.
 XX PR 22-OCT-2001; 2001EP-0000558.
 XX PR 15-NOV-2001; 2001EP-0004464.
 XX PR 21-DEC-2001; 2001EP-0005117.
 XX PA (STAM) DSM NV.
 XX PI Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 XX PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 XX PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 XX PI Stiebler J, Albarg R;
 XX DR WPI; 2002-723203/78.
 XX DR P-PSDB; AB278231, AB278288.
 XX PT Novel isolated protease polypeptide useful in laboratory, clinical,
 XX PT pharmaceutical, chemical, diagnostic, personal care and industrial
 XX PT applications -
 XX PS Claim 13; Page 379-382; 394pp; English.
 XX CC The invention relates to a novel isolated protease polypeptide. A
 XX CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 XX CC fungal infection such as aspergillosis, or as a query sequence to perform

a search against public databases. A polypeptide of the invention is
 useful in a selected number of industrial or pharmaceutical processes, in
 laboratory or clinical processes, in food industry (baking, brewing,
 cheese manufacturing, meat tenderising), in tanning industry and in the
 manufacture of biological detergents. A polypeptide may also be useful
 for improving protein solubility, extraction yields, viscosity or taste,
 texture, nutritional value, minimising of antigenicity or
 anti-nutritional factors, colour or functionality as well as processing
 aspects like filterability of the proteinaceous raw material. The
 sequences shown in ABR38913-ABR38869 represent the A. niger proteases of
 the invention.

XX SQ Sequence 551 AA;
 Query Match 4.3%; Score 24; DB 23; Length 551;
 Best Local Similarity 100.0%; Pred. No. 3.6e-15;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPGSDSLIGLFEELGPCH 122
 |||||
 Db 99 TLWLNGGPGSDSLIGLFEELGPCH 122

RESULT 4
 AAW56100
 ID AAW56100 standard; Protein; 20 AA.
 XX AC AAW56100;
 XX DT 27-AUG-1998 (first entry)
 DE A. oryzae ATCC20386 carboxypeptidase I peptide fragment #1.
 XX Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
 KW food industry.
 XX OS Aspergillus oryzae.
 XX PH Key Location/Qualifiers
 FT Modified-site 1
 FT /label= unknown
 XX WO9814599-A1.
 XX PD 09-APR-1998.
 XX PF 03-OCT-1997; 97WO-US17977.
 XX PR 27-NOV-1996; 96US-0757534.
 XX PR 04-OCT-1996; 96US-0726880.
 XX PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX PI Berka R, Blinkovsky A, Brown K, Danbmann C, Golightly E;
 XX PI Klotz A, Mathisen TE, Rey M;
 XX DR WPI; 1998-240098/21.
 XX PT Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 XX PT substrates, useful for improving flavour of foods
 XX PS Example 2; Page 41; 82pp; English.
 XX CC This sequence represents an N-terminal peptide from carboxypeptidase I
 CC isolated from Aspergillus oryzae. This polypeptide has an optimal
 CC activity in the range of pH 3.0-7.5 at 25 deg. C, optimal activity in
 CC the range of 55-60 deg. C at pH 4, and a residual activity of at least
 CC 65.5% after 30 minutes at pH 4.0 and 60 deg. C. It also has the capacity
 CC to hydrolyse X from N-CBZ-Ala-X where N-CBZ is N-carboxy and X is
 CC any amino acid. The carboxypeptidases can be used for obtaining
 CC hydrolysates (which can be enriched in free glutamic acid or peptide
 CC bound glutamic acid residues) from proteinaceous substrates. The

CC carboxypeptidases can be used in flavour-improving compositions in the
CC food industry. The products can also be used for the production of
CC polypeptides free of carboxypeptidase activity.

XX SQ Sequence 20 AA;

Query Match 3.4%; Score 19; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 YGCHYGPAFFNHFYQNER 248
DB 2 YGCHYGPAFFNHFYQNER 20
|||||

RESULT 5

AAW57041
ID AAW57041 standard; Protein; 29 AA.

XX AC AAW57041;

XX DT 27-AUG-1998 (first entry)

XX DE A. oryzae ATCC20386 carboxypeptidase I peptide fragment #4.

XX KW Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
XX food industry.

XX OS Aspergillus oryzae.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT Modified-site 10 /label= unknown

FT Modified-site 10 /label= unknown

XX WO9814599-A1.

XX PD 09-APR-1998.

XX PF 03-OCT-1997; 97WO-US17977.

XX PR 27-NOV-1996; 96US-0757534.

XX PR 04-OCT-1996; 96US-0726880.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;

XX PI Klotz A, Mathisen TE, Rey M;

XX XX WPI; 1998-240098/21.

XX PT Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
XX PT substrates, useful for improving flavour of foods

XX PS Example 2; Page 42; 82pp; English.

XX CC This sequence represents a peptide fragment from carboxypeptidase I
XX CC isolated from Aspergillus oryzae and generated from cyanogen bromide
XX CC cleavage. This polypeptide has an optimal activity in the range of pH
XX CC 3.0-7.5 at 25 deg. C, optimal activity in the range of 55-60 deg. C at
XX CC pH 4, and a residual activity of at least 65.5% after 30 minutes at pH
XX CC 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from
XX CC N-CBZ-Ala-X where N-CBZ is N-carboxybenzoyl and X is any amino acid. The
XX CC carboxypeptidases can be used for obtaining hydrolysates (which can be
XX CC enriched in free glutamic acid or peptide bound glutamic acid residues)
XX CC from proteinaceous substrates. The carboxypeptidases can be used in
XX CC flavour-improving compositions in the food industry. The products can
XX CC also be used for the production of polypeptides free of carboxypeptidase
XX CC activity.

XX SQ Sequence 29 AA;

Query Match 2.9%; Score 16; DB 19; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 YAFAGRGVYDIRHPYD 364
DB 11 YAFAGRGVYDIRHPYD 26
|||||

RESULT 6

ABR38817
ID ABR38817 standard; Protein; 526 AA.

XX AC ABR38817;

XX DT 24-APR-2003 (first entry)

XX DE A. niger serine carboxypeptidase polypeptide #1.

XX KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
XX KW protein solubility; viscosity; taste; texture; nutritional value;
XX KW EC3.4.16.6.

XX OS Aspergillus niger.

XX PN WO200268623-A2.

XX PD 06-SEP-2002.

XX PF 22-FEB-2002; 2002WO-EP01984.

XX PR 23-FEB-2001; 2001EP-0200657.

XX PR 23-FEB-2001; 2001EP-0200658.

XX PR 23-FEB-2001; 2001EP-0200660.

XX PR 26-FEB-2001; 2001EP-0200706.

XX PR 26-FEB-2001; 2001EP-0200707.

XX PR 26-FEB-2001; 2001EP-0200708.

XX PR 26-FEB-2001; 2001EP-0200719.

XX PR 28-MAR-2001; 2001EP-0000075.

XX PR 28-MAR-2001; 2001EP-0000078.

XX PR 28-MAR-2001; 2001EP-0000080.

XX PR 28-MAR-2001; 2001EP-0000087.

XX PR 28-MAR-2001; 2001EP-0000088.

XX PR 21-MAY-2001; 2001EP-0000156.

XX PR 21-MAY-2001; 2001EP-0000159.

XX PR 21-MAY-2001; 2001EP-0000160.

XX PR 21-MAY-2001; 2001EP-0000162.

XX PR 21-MAY-2001; 2001EP-0000165.

XX PR 21-MAY-2001; 2001EP-0000166.

XX PR 21-MAY-2001; 2001EP-0000168.

XX PR 21-JUN-2001; 2001EP-0000240.

XX PR 21-JUN-2001; 2001EP-0000242.

XX PR 21-JUN-2001; 2001EP-0000244.

XX PR 21-JUN-2001; 2001EP-0000246.

XX PR 12-JUL-2001; 2001EP-0000280.

XX PR 12-JUL-2001; 2001EP-0000285.

XX PR 30-JUL-2001; 2001EP-0000323.

XX PR 30-JUL-2001; 2001EP-0000327.

XX PR 02-AUG-2001; 2001EP-0000341.

XX PR 02-AUG-2001; 2001EP-0000342.

XX PR 02-AUG-2001; 2001EP-0000343.

XX PR 02-AUG-2001; 2001EP-0000344.

XX PR 09-AUG-2001; 2001EP-0000357.

XX PR 16-AUG-2001; 2001EP-0000374.

XX PR 16-AUG-2001; 2001EP-0000377.

XX PR 20-SEP-2001; 2001EP-0000478.

XX PR 20-SEP-2001; 2001EP-0000483.

XX PR 22-OCT-2001; 2001EP-0000552.

XX PR 22-OCT-2001; 2001EP-0000553.

XX PR 22-OCT-2001; 2001EP-0000554.

XX PR 22-OCT-2001; 2001EP-0000556.

XX PR 22-OCT-2001; 2001EP-0000557.

```

PR 22-OCT-2001; 2001EP-0000558.
PR 15-NOV-2001; 2001EP-0004464.
PR 21-DEC-2001; 2001EP-0005117.
XX (STAM) DSM NV.
PA
XX
XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
PI Stiebler J, Albang R;
XX
XX WPI; 2002-723203/78.
DR P-PSDB; ABZ78184, ABZ78241.
XX
XX Novel isolated protease polypeptide useful in laboratory, clinical,
PT pharmaceutical, chemical, diagnostic, personal care and industrial
PT applications
XX
XX Claim 13; Page 271-274; 394pp; English.
XX
XX The invention relates to a novel isolated protease polypeptide. A
CC polypeptide or polynucleotide of the invention is useful for diagnosing a
CC fungal infection such as aspergillosis, or as a query sequence to perform
CC a search against public databases. A polypeptide of the invention is
CC useful in a selected number of industrial or pharmaceutical processes, in
CC laboratory or clinical processes, in food industry (baking, brewing,
CC cheese manufacturing, meat tenderizing), in tanning industry and in the
CC manufacture of biological detergents. A polypeptide may also be useful
CC for improving protein solubility, extraction yields, viscosity or taste,
CC texture, nutritional value, minimising of antigenicity or
CC anti-nutritional factors, colour or functionality as well as processing
CC aspects like filterability of the proteinaceous raw material. The
CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
CC the invention.
XX
XX SQ Sequence 526 AA;

Query Match 1.8%; Score 10; DB 23; Length 526;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TSYGGHYGP 236
Db 181 TSYGGHYGP 190
|||||

RESULT 7
AAU46801
ID AAU46801 standard; Protein; 178 AA.
AC
XX AAU46801;
XX
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #7697.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypervitosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX
XX WO200181581-A2.
PN
XX
XX 01-NOV-2001.
PD
XX
XX 20-APR-2001; 2001WO-US12865.
PF
XX
XX 21-APR-2000; 2000US-199047P.
PR
XX 02-JUN-2000; 2000US-208841P.
PR
XX 07-JUL-2000; 2000US-216747P.
PR
XX

(CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59535.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 7996; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypervitosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA)
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 178 AA;

Query Match 1.6%; Score 9; DB 22; Length 178;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FLSVLPLVA 14
Db 151 FLSVLPLVA 159
|||||

RESULT 8
AAG17034
ID AAG17034 standard; Protein; 454 AA.
AC
XX AAG17034;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 17902.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX 05-MAR-1999; 99US-0123180.
PR
XX 09-MAR-1999; 99US-0123548.
PR
XX 23-MAR-1999; 99US-0125788.
PR
XX 25-MAR-1999; 99US-0126264.

```

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132566.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159638.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160988.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 9; DB 21; Length 454;
 Best Local Similarity 100.0%; Pred. No. 4.8;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 TLWLNGGPG 107

Db 59 TLWLNGGPG 67

RESULT 9

AAG17033
 ID AAG17033 standard; Protein; 461 AA.

XX AC AAG17033;

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 17901.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0134941.
 PR 21-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137582.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149388.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156459.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 9; DB 21; Length 461;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPG 107
|||||||
Db 66 TLWLNGGPG 74

RESULT 10
AAG49759
ID AAG49759 standard; Protein; 461 AA.
XX
AC AAG49759;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62381.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

```
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match          1.6%; Score 9; DB 21; Length 461;
Best Local Similarity 100.0%; Pred. NO. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPG 107
DB 66 TLWLNGGPG 74

RESULT 11
AAG49762
ID AAG49762 standard; Protein; 461 AA.
XX AC AAG49762;
XX XX
DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 62985.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX XX
XX PN EP1033405-A2.
XX XX
XX PD 06-SEP-2000.
XX XX
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 26-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 18-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144894.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
```


PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159233.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 9; DB 21; Length 461;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPG 107
|||
DB 66 TLWLNGGEG 74

RESULT 12

AAAG17032
ID AAAG17032 standard; Protein; 473 AA.

XX
AC AAAG17032;

XX
DT 17-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17900.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
EF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

us-09-712-338-2.oli.rag

Mon Nov 24 13:41:05 2003

PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139459.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149226.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0140353.	PR	30-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142254.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144333.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159638.
PR	20-JUL-1999;	99US-0144632.	PR	18-OCT-1999;	99US-0159584.
PR	20-JUL-1999;	99US-0144884.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160770.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160815.
PR	22-JUL-1999;	99US-0145089.	PR	22-OCT-1999;	99US-0160980.
PR	22-JUL-1999;	99US-0145192.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160989.
PR	23-JUL-1999;	99US-0145218.	PR	25-OCT-1999;	99US-0161404.
PR	23-JUL-1999;	99US-0145224.	PR	25-OCT-1999;	99US-0161405.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145911.	PR	25-OCT-1999;	99US-0161359.
PR	27-JUL-1999;	99US-0145918.	PR	26-OCT-1999;	99US-0161360.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161361.
PR	28-JUL-1999;	99US-0145951.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146386.	PR	28-OCT-1999;	99US-0161992.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161992.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161993.
PR	03-AUG-1999;	99US-0147038.			

```
PR 29-OCT-1999; 99US-0162142.
Query Match 1.6%; Score 9; DB 21; Length 473;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 99 TLWLNGGPG 107
Db 78 TLWLNGGPG 86

RESULT 13
AAG49758
ID AAG49758 standard; Protein; 473 AA.
XX
AC AAG49758;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62980.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123588.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
```

09-AUG-1999;	99US-0147935.	DB		78 TLWNGGPG 86
10-AUG-1999;	99US-0148171.			
11-AUG-1999;	99US-0148319.			
12-AUG-1999;	99US-0148341.			
13-AUG-1999;	99US-0148565.			
13-AUG-1999;	99US-0148684.			
16-AUG-1999;	99US-0149368.			
17-AUG-1999;	99US-0149175.			
18-AUG-1999;	99US-0149426.			
20-AUG-1999;	99US-0149722.			
20-AUG-1999;	99US-0149723.			
20-AUG-1999;	99US-0149929.			
23-AUG-1999;	99US-0149902.			
23-AUG-1999;	99US-0149930.			
25-AUG-1999;	99US-0150566.			
26-AUG-1999;	99US-0150884.			
27-AUG-1999;	99US-0151065.			
27-AUG-1999;	99US-0151066.			
27-AUG-1999;	99US-0151080.			
30-AUG-1999;	99US-0151303.			
31-AUG-1999;	99US-0151438.			
01-SEP-1999;	99US-0151930.			
07-SEP-1999;	99US-0152363.			
10-SEP-1999;	99US-0153070.			
13-SEP-1999;	99US-0153758.			
15-SEP-1999;	99US-0154018.			
16-SEP-1999;	99US-0154039.			
20-SEP-1999;	99US-0154779.			
22-SEP-1999;	99US-0155139.			
23-SEP-1999;	99US-0155486.			
24-SEP-1999;	99US-0155659.			
28-SEP-1999;	99US-0156458.			
29-SEP-1999;	99US-0156596.			
04-OCT-1999;	99US-0157117.			
05-OCT-1999;	99US-0157753.			
06-OCT-1999;	99US-0157865.			
07-OCT-1999;	99US-0158029.			
08-OCT-1999;	99US-0158232.			
12-OCT-1999;	99US-0158369.			
13-OCT-1999;	99US-0159293.			
13-OCT-1999;	99US-0159294.			
13-OCT-1999;	99US-0159295.			
14-OCT-1999;	99US-0159329.			
14-OCT-1999;	99US-0159330.			
14-OCT-1999;	99US-0159331.			
14-OCT-1999;	99US-0159637.			
14-OCT-1999;	99US-0159638.			
18-OCT-1999;	99US-0159584.			
21-OCT-1999;	99US-0160741.			
21-OCT-1999;	99US-0160767.			
21-OCT-1999;	99US-0160768.			
21-OCT-1999;	99US-0160770.			
21-OCT-1999;	99US-0160814.			
21-OCT-1999;	99US-0160815.			
22-OCT-1999;	99US-0160980.			
22-OCT-1999;	99US-0160981.			
22-OCT-1999;	99US-0160989.			
25-OCT-1999;	99US-0161404.			
25-OCT-1999;	99US-0161405.			
25-OCT-1999;	99US-0161406.			
26-OCT-1999;	99US-0161359.			
26-OCT-1999;	99US-0161360.			
26-OCT-1999;	99US-0161361.			
28-OCT-1999;	99US-0161920.			
28-OCT-1999;	99US-0161992.			
28-OCT-1999;	99US-0161993.			
29-OCT-1999;	99US-0162144.			
Query Match 1.6%; Score 9; DB 21; Length 473;				
Best Local Similarity 100.0%; Pred.No. 4.9;				
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
99 TLWNGGPG 107				
Qy				
DB	78 TLWNGGPG 86			
RESULT 14				
AG49761				
ID	AAAG49761 standard; Protein; 473 AA.			
AC	AAAG49761;			
XX	18-OCT-2000 (first entry)			
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 62984.			
XX	Protein identification; signal transduction pathway; metabolic pathway;			
DE	KW hybridisation assay; genetic mapping; gene expression control; promoter;			
XX	termination sequence.			
OS	Arabidopsis thaliana.			
XX	PF1033405-A2.			
PN	06-SEP-2000.			
PD	25-FEB-2000; 2000EP-0301439.			
XX	25-FEB-1999;	99US-0121825.		
PR	05-MAR-1999;	99US-0123180.		
PR	09-MAR-1999;	99US-0123548.		
PR	23-MAR-1999;	99US-0125788.		
PR	25-MAR-1999;	99US-0126264.		
PR	29-MAR-1999;	99US-0126785.		
PR	01-APR-1999;	99US-0127462.		
PR	06-APR-1999;	99US-0128234.		
PR	08-APR-1999;	99US-0128714.		
PR	16-APR-1999;	99US-0129845.		
PR	19-APR-1999;	99US-0130077.		
PR	21-APR-1999;	99US-0130449.		
PR	23-APR-1999;	99US-0130510.		
PR	23-APR-1999;	99US-0130891.		
PR	28-APR-1999;	99US-0131449.		
PR	30-APR-1999;	99US-0132048.		
PR	04-MAY-1999;	99US-0132407.		
PR	05-MAY-1999;	99US-0132484.		
PR	06-MAY-1999;	99US-0132485.		
PR	06-MAY-1999;	99US-0132486.		
PR	07-MAY-1999;	99US-0132487.		
PR	11-MAY-1999;	99US-0132863.		
PR	14-MAY-1999;	99US-0134256.		
PR	14-MAY-1999;	99US-0134218.		
PR	14-MAY-1999;	99US-0134219.		
PR	14-MAY-1999;	99US-0134221.		
PR	14-MAY-1999;	99US-0134370.		
PR	18-MAY-1999;	99US-0134768.		
PR	19-MAY-1999;	99US-0134941.		
PR	20-MAY-1999;	99US-0135124.		
PR	21-MAY-1999;	99US-0135353.		
PR	24-MAY-1999;	99US-0135629.		
PR	25-MAY-1999;	99US-0136021.		
PR	27-MAY-1999;	99US-0136392.		
PR	28-MAY-1999;	99US-0136782.		
PR	01-JUN-1999;	99US-0137222.		
PR	03-JUN-1999;	99US-0137528.		
PR	04-JUN-1999;	99US-0137502.		
PR	07-JUN-1999;	99US-0137724.		
PR	08-JUN-1999;	99US-0138094.		
PR	10-JUN-1999;	99US-0138540.		
PR	10-JUN-1999;	99US-0138847.		
PR	14-JUN-1999;	99US-0139119.		
PR	16-JUN-1999;	99US-0139452.		
PR	16-JUN-1999;	99US-0139453.		
PR	17-JUN-1999;	99US-0139492.		
PR	18-JUN-1999;	99US-0139454.		

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 9; DB 21; Length 473;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGPG 107
|||
Db 78 TLWLNGPG 86
|||

RESULT 15
ABG93049
ID ABG93049 standard; Protein; 508 AA.

•